

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2005, 23:17:42 : Search time 101 Seconds
(without alignments)
3127.531 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWMAQEREDIIIM.....TSGPCLCGHLESQPPPGF 549

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	549	1	TBC3 HUMAN
2	2971	99.6	549	2	Q61PXL
3	2971	99.6	549	2	AARH1680
4	2945	98.7	549	2	Q6DHY5
5	2188.5	73.4	1406	1	UBPE_HUMAN
6	1544.5	51.8	376	2	Q15635
7	1479	49.6	346	2	Q6P722
8	1479	49.6	346	2	AAH58890
9	1477.5	49.5	376	2	Q6P1J2
10	1477.5	49.5	376	2	AAH33670
11	1346	45.1	291	1	Q6DCB4
12	655	22.0	819	1	USGL_MOUSE
13	655	22.0	819	1	BAC97847
14	651	21.8	828	1	USGL_HUMAN
15	651	21.8	838	2	BAAO2807
16	603.5	20.2	250	2	Q86UD7
17	602	20.2	224	2	Q8NCR5
18	576	19.3	1085	2	Q9U2D8
19	576	19.3	1085	2	CAB60374
20	574.5	19.3	1021	2	Q7K711
21	574.5	19.3	1021	2	CAE5741
22	530	17.8	391	2	Q7QCV5
23	500.5	16.3	498	2	Q7PEV8
24	455	15.3	457	2	Q18357
25	455	15.3	485	2	Q9SRE0
26	455	15.3	571	2	Q7KR75
27	323.5	10.8	356	2	Q81347
28	321.5	10.8	537	2	Q8BHL3
29	321.5	10.8	537	2	Q91XR3
30	321.5	10.8	655	2	Q6CGW9
31	320.5	10.7	533	2	Q9H822

32	320.5	10.7	584	2	Q6IN54	Q6IN54 homo sapien
33	320.5	10.7	584	2	AARH2453	AARH2453 homo sapi
34	320.5	10.7	622	2	Q6P530	Q6P530 homo sapien
35	320.5	10.7	622	2	AAH63112	AAH63112 homo sapi
36	319	10.7	353	2	Q8LAF3	Q8LAF3 arabidopsis
37	317.5	10.6	304	2	Q9M894	Q9M894 arabidopsis
38	306.5	10.3	532	2	Q9Y1V8	Q9Y1V8 halocynthia
39	305.5	10.2	500	1	TB10_MOUSE	P58802 mus muscucu
40	302	10.1	298	2	Q8L756	Q8L756 arabidopsis
41	301	10.1	446	2	Q8IV04	Q8IV04 homo sapien
42	301	10.1	446	2	AAH62999	AAH62999 homo sapi
43	301	10.1	450	2	AAH36873	AAH36873 homo sapi
44	300	10.1	500	2	Q8CI62	Q8CI62 mus muscucu
45	299.5	10.0	508	1	TB10_HUMAN	Q9BX16 homo sapien

ALIGNMENTS

RESULT 1
TBC3_HUMAN STANDARD; PRT; 549 AA.
ID TBC3_HUMAN
AC Q81ZP1; QSH0B9; Q9UDD4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE TBC1 domain family member 3 (Rab GTPase-activating protein PRG17)
DE (Prostate cancer gene 17 protein) (TRB17 alpha protein).
GN Name=TBC1D3; Synonyms=PRG17;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, AND DISEASE.
RX MEDLINE=22246263; PubMed=12359748;
RA Pei L., Peng Y., Yang Y., Ling X.B., Van Eynhoven W.G., Nguyen K.C.,
RA Rublin M., Hoey T., Powers S., Li J.;
RT "PRG17, a novel oncogene encoding a Rab GTPase-activating protein, is
RT amplified in prostate cancer.";
RL Cancer Res. 62:5420-5424(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RX MEDLINE=21154917; PubMed=11230166; DOI=10.1101/gr.154701;
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glasel S.,
RA Ansgore W., Beecher M., Bloeker H., Bauesche S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Olterswaelder B., Obermaier B., Tampe J., Heubner D.,
RA Twardant R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [3]
RP SEQUENCE OF 1-52 FROM N.A.
RX MEDLINE=9328825; PubMed=8471161;
RA Onno M., Nakamura T., Mariga-Samson R., Hillova J., Hill M.;
RT "Human TRB17 oncogene is generated from a family of homologous
RT polynucleotide sequences by single-base changes.";
RL DNA Cell Biol. 12:107-118(1993).
RN [4]
RP SEQUENCE OF 1-52 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94010310; PubMed=8406013;
RA Onno M., Nakamura T., Hillova J., Hill M.;
RT "Identification of novel sequences in the repertoire of hypervariable
RT TRB17 genes from immortalized normal and malignant human
RT keratinocytes.";
RL Gene 131:209-215(1993).
RN [5]
RP TISSUE SPECIFICITY.
RX MEDLINE=22506395; PubMed=12604796; DOI=10.1073/pnas.0437015100;
RA Paulding C.A., Ruvoletto M., Haber D.A.;

RT "The Trel2 (USP6) oncogene is a hominoid-specific gene."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:2507-2511 (2003).
 CC -1- FUNCTION: Acts as a GTPase activating protein for RAB5. Does not
 CC act on RAB4 or RAB11.
 CC -1- TISSUE SPECIFICITY: Expressed in liver, skeletal muscle, kidney,
 CC pancreas, spleen, testis, ovary, small intestine and peripheral
 CC blood leukocytes. Overexpressed in prostate cancers.
 CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----
 DR EMBL; AF540953; AAN3117.1; -;
 DR EMBL; AL136860; CAB66794.1; -;
 DR EMBL; X71377; CAB94197.1; -;
 DR Genew; HGNC:19031; TBCID3.
 DR MIM; 607741; -;
 DR InterPro; IPR000195; RabGAP_TBC.
 DR Pfam; PF00566; TBC; 1.
 DR SMART; SM00164; TBC; 1.
 DR PROSITE; PS50086; TBC_RABGAP; 1.
 DR KEGG; K05008; TBC_RABGAP; 1.
 DR GTPase activation; Proto-oncogene.
 FT DOMAIN 101 293
 FT CONFLICT 117 117 Rab-GAP TBC.
 FT CONFLICT 137 137 I -> T (in Ref. 2).
 FT CONFLICT 137 137 R -> K (in Ref. 2).
 FT CONFLICT 354 354 Q -> K (in Ref. 2).
 FT CONFLICT 526 526 P -> Q (in Ref. 2).
 SQ SEQUENCE 549 AA; 62168 MW; 831DA747DOB6400 CRC64;

Query Match 100.0%; Score 2983; DB 1; Length 549;
 Best Local Similarity 100.0%; Pred. No. 1.4e-179;
 Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDVEVAGSMAOEREDIMKYEKGHRAGLPEDGKPPRSYNNVDHIGIVHETELPPL 60
 DB 1 MDVEVAGSMAOEREDIMKYEKGHRAGLPEDGKPPRSYNNVDHIGIVHETELPPL 60
 QY 61 TAREAKQIRREISRSKSKVMDLGDMEKYSRKLIDRAYKGMPIIRGPMVSLNIEEM 120
 DB 61 TAREAKQIRREISRSKSKVMDLGDMEKYSRKLIDRAYKGMPIIRGPMVSLNIEEM 120
 QY 121 KLNKPGRYQIMKEKGRSSHHIQRIDRVSGTLRKHIIFRDYGTQKRELLHILAYEY 180
 DB 121 KLNKPGRYQIMKEKGRSSHHIQRIDRVSGTLRKHIIFRDYGTQKRELLHILAYEY 180
 QY 181 NPEVGYRDLSTHIALFLYLPEDEAFWALVOLLAERHSLQGRHSPNGVQGLDQOE 240
 DB 181 NPEVGYRDLSTHIALFLYLPEDEAFWALVOLLAERHSLQGRHSPNGVQGLDQOE 240
 QY 241 HVAATSOPTKMGHODKDLGQCSPLGLIRLLIDGSLGTLRLMDVYLVEBQALMP 300
 DB 241 HVAATSOPTKMGHODKDLGQCSPLGLIRLLIDGSLGTLRLMDVYLVEBQALMP 300
 QY 301 TRIAFVQOQRLTKTSRCGPWAFPCNRFVDTWARDEDTYVAKHRAASKULTRKQGLPP 360
 DB 301 TRIAFVQOQRLTKTSRCGPWAFPCNRFVDTWARDEDTYVAKHRAASKULTRKQGLPP 360
 QY 361 APEEGSSASRPVPSRSGKTLCKGDQAPGPPAPRPPIISASPPRASPSTCPCGA 420
 DB 361 APEEGSSASRPVPSRSGKTLCKGDQAPGPPAPRPPIISASPPRASPSTCPCGA 420
 QY 421 VREDIYPVGTGVSPPLAQCGPGSWMFLQNNMSPRLPTDLDVGPMPRRHYDPRQSCW 480
 DB 421 VREDIYPVGTGVSPPLAQCGPGSWMFLQNNMSPRLPTDLDVGPMPRRHYDPRQSCW 480
 QY 481 RAISOEDQAPCWAHEPAREVRSFAAPSTSDQCTPPRARDPCAPTSGPCLCGIHL 540
 DB 481 RAISOEDQAPCWAHEPAREVRSFAAPSTSDQCTPPRARDPCAPTSGPCLCGIHL 540

DB 481 RAISOEDQAPCWAHEPAREVRSFAAPSTSDQCTPPRARDPCAPTSGPCLCGIHL 540
 QY 541 ESSQPPPGF 549
 DB 541 ESSQPPPGF 549

RESULT 2

ID OGIPI1 PRELIMINARY; PRT; 549 AA.
 AC OGIPI1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 GN TBC1 domain family, member 3.
 GN Name=TBCID3;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marisina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stopleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueding T.B., Toshiyuki S., Carninci P., Mulhaby S.J.,
 RA Rana S.S., Loggani N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Murray D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzyzanski M.T., Skalska V., Smalius D.B., Schmechel A., Schein J.E.,
 RA Jones S.J., Maira M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX Strausberg R.L.;
 RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC071680; AAH71680.1; -;
 DR InterPro; IPR000195; RabGAP_TBC.
 DR Pfam; PF00566; TBC; 1.
 DR SMART; SM00164; TBC; 1.
 DR PROSITE; PS50086; TBC_RABGAP; 1.
 SQ SEQUENCE 549 AA; 62231 MW; 5657B5602B3655FS CRC64;

Query Match 99.6%; Score 2971; DB 2; Length 549;
 Best Local Similarity 99.5%; Pred. No. 8.2e-179;
 Matches 546; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVEVAGSMAOEREDIMKYEKGHRAGLPEDGKPPRSYNNVDHIGIVHETELPPL 60
 DB 1 MDVEVAGSMAOEREDIMKYEKGHRAGLPEDGKPPRSYNNVDHIGIVHETELPPL 60
 QY 61 TAREAKQIRREISRSKSKVMDLGDMEKYSRKLIDRAYKGMPIIRGPMVSLNIEEM 120
 DB 61 TAREAKQIRREISRSKSKVMDLGDMEKYSRKLIDRAYKGMPIIRGPMVSLNIEEM 120
 QY 121 KLNKPGRYQIMKEKGRSSHHIQRIDRVSGTLRKHIIFRDYGTQKRELLHILAYEY 180
 DB 121 KLNKPGRYQIMKEKGRSSHHIQRIDRVSGTLRKHIIFRDYGTQKRELLHILAYEY 180
 QY 181 NPEVGYRDLSTHIALFLYLPEDEAFWALVOLLAERHSLQGRHSPNGVQGLDQOE 240
 DB 181 NPEVGYRDLSTHIALFLYLPEDEAFWALVOLLAERHSLQGRHSPNGVQGLDQOE 240

```

DB 181 NBEVGCRLSHIALFLILYLPBEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240
QY 241 HVAATSOPKTMGHODKKDLCCGCSPLGCLIRLIDGISLGLTRLMDVYLVEGEOLMPT 300
DB 241 HVAATSOPKTMGHODKKDLCCGCSPLGCLIRLIDGISLGLTRLMDVYLVEGEOLMPT 300
QY 301 TRIAFVQOQRLTKTSRCGPMARFCNRPVDTWABEDTVLKHLRASMKKLTRKQGLPPP 360
DB 301 TRIAFVQOQRLTKTSRCGPMARFCNRPVDTWABEDTVLKHLRASMKKLTRKQGLPPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDROAPGPPAPRPPIWASPPRASPSTPCPGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDROAPGPPAPRPPIWASPPRASPSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRLQWNSMRLPTDLVGEPMFRHYDFQSCV 480
DB 421 VREDTYPVGTQGVSPALAOGGPQGSWRLQWNSMRLPTDLVGEPMFRHYDFQSCV 480
QY 481 RAISOEDOLAPCQOAEHPAERVSAPAPSTDSOGTFFRARDQPCAPTSGPCLCGLHL 540
DB 481 RAISOEDOLAPCQOAEHPAERVSAPAPSTDSOGTFFRARDQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPPGF 549
DB 541 ESSQFPPGF 549

RESULT 3
AAH71680 PRELIMINARY; PRT; 549 AA.
ID AAH71680;
AC AAH71680;
DT 01-JUN-2004 (TReMBLrel. 27, Created)
DT 01-JUN-2004 (TReMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TReMBLrel. 27, Last annotation update)
DE TBC1 domain family, member 3.
GN TBC1D3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Brain;
RC MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stachleon M., Soares M.B., Bonaldo M.F., Caeavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP TISSUE=Brain;
RC SEQUENCE FROM N.A.
RA Strubeberg R.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071680; AAH71680.1; -
RA SEQUENCE 549 AA; 62231 MW; 5657E5602B3655F5 CRC64;

```

```

Query Match 99.6%; Score 2971; DB 2; Length 549;
Best Local Similarity 99.5%; Pred. No. 8 2e-179;
Matches 546; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVVEAGSWMAQOEEDIIIMKEKGRAGLPEDKPKPPRSYNNNVHIGIVHETLPP 60
DB 1 MDVVEAGSWMAQOEEDIIIMKEKGRAGLPEDKPKPPRSYNNNVHIGIVHETLPP 60
QY 61 TAREAKQIRREISRSKSWMDLGDWEKYSKSLIDRAYKGMPMNIRGPMVLLNIEEM 120
DB 61 TAREAKQIRREISRSKSWMDLGDWEKYSKSLIDRAYKGMPMNIRGPMVLLNIEEM 120
QY 121 KLNKRGVQIMKEKGRSEHIIQIDRVSGTLRGHIFPRDRYGTQKORELHILAYEY 180
DB 121 KLNKRGVQIMKEKGRSEHIIQIDRVSGTLRGHIFPRDRYGTQKORELHILAYEY 180
QY 181 NBEVGCRLSHIALFLILYLPBEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240
DB 181 NBEVGCRLSHIALFLILYLPBEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240
QY 241 HVAATSOPKTMGHODKKDLCCGCSPLGCLIRLIDGISLGLTRLMDVYLVEGEOLMPT 300
DB 241 HVAATSOPKTMGHODKKDLCCGCSPLGCLIRLIDGISLGLTRLMDVYLVEGEOLMPT 300
QY 301 TRIAFVQOQRLTKTSRCGPMARFCNRPVDTWABEDTVLKHLRASMKKLTRKQGLPPP 360
DB 301 TRIAFVQOQRLTKTSRCGPMARFCNRPVDTWABEDTVLKHLRASMKKLTRKQGLPPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDROAPGPPAPRPPIWASPPRASPSTPCPGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDROAPGPPAPRPPIWASPPRASPSTPCPGA 420
QY 421 VREDTYPVGTQGVSPALAOGGPQGSWRLQWNSMRLPTDLVGEPMFRHYDFQSCV 480
DB 421 VREDTYPVGTQGVSPALAOGGPQGSWRLQWNSMRLPTDLVGEPMFRHYDFQSCV 480
QY 481 RAISOEDOLAPCQOAEHPAERVSAPAPSTDSOGTFFRARDQPCAPTSGPCLCGLHL 540
DB 481 RAISOEDOLAPCQOAEHPAERVSAPAPSTDSOGTFFRARDQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPPGF 549
DB 541 ESSQFPPGF 549

RESULT 4
Q6DHY5 PRELIMINARY; PRT; 549 AA.
ID Q6DHY5;
AC Q6DHY5;
DT 01-OCT-2004 (TReMBLrel. 28, Created)
DT 01-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP TISSUE=Brain;
RC MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stachleon M., Soares M.B., Bonaldo M.F., Caeavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,

```

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Myers R.M., Butterfield J.S.,
 RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerh A., Schein J.E.,
 RA Jones S.J., Maita W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC075809; AAT75809.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 549 AA; 62304 MW; 401430A45C00D9C5 CRC64;

Query Match 98.7%; Score 2945; DB 2; Length 549;
 Best Local Similarity 98.5%; Pred. No. 3.5e-177;
 Matches 541; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MDVEVAGSWMAQREDIMKYEKGKRAGLPEDEKPKPFRSYNNNDHIGIVHETLPL 60
 DB 1 MDVEVAGSWMAQREDIMKYEKGKRAGLPEDEKPKPFRSYNNNDHIGIVHETLPL 60
 QY 61 TAPBAKQIRREIRSKSKVMDLGDWEKYKSKRLIDRAYKGMPTIRGPMWSVLNIEM 120
 DB 61 TAPBAKQIRREIRSKSKVMDLGDWEKYKSKRLIDRAYKGMPTIRGPMWSVLNIEM 120
 QY 121 KLANPGYQIMKEKGRSSSHIQRIDRVSGTRKHIFPDRDGTQORLHILAYEY 180
 DB 121 KLANPGYQIMKEKGRSSSHIQRIDRVSGTRKHIFPDRDGTQORLHILAYEY 180
 QY 181 NPEVGCRLDLSHIAALFLYLPEEDAFAVLVOLLASERHSLQGFHSPNGTVOGLDQOE 240
 DB 181 NPEVGCRLDLSHIAALFLYLPEEDAFAVLVOLLASERHSLQGFHSPNGTVOGLDQOE 240
 QY 241 HVAATQPKMGHODKDDCGGSPICILIRLIDGISTGLTRLMDVTVLVEGQALMP 300
 DB 241 HVAATQPKMGHODKDDCGGSPICILIRLIDGISTGLTRLMDVTVLVEGQALMP 300
 QY 301 TRIAFVQOQRLTKTSRCGPWFARFCNRFVDTWARDEDVTLKHLRASMKKLTPKQODLPP 360
 DB 301 TRIAFVQOQRLTKTSRCGPWFARFCNRFVDTWARDEDVTLKHLRASMKKLTPKQODLPP 360
 QY 361 APEQSSASRPVPSRSGKTKCKGDRQAPPEPPARPPPIWSASPPRAPSSTPCPGA 420
 DB 361 APEQSSASRPVPSRSGKTKCKGDRQAPPEPPARPPPIWSASPPRAPSSTPCPGA 420
 QY 421 VAEEDTVPGTQVSPALAQGGPQGSWFLQNSMPLPTDLDVYGFPFRHYDRQSCWV 480
 DB 421 VAEEDTVPGTQVSPALAQGGPQGSWFLQNSMPLPTDLDVYGFPFRHYDRQSCWV 480
 QY 481 RAISEDOLAPCQWQEHFAERVSFAAPSTDSQGTFRARDQPCATSGPLCLGHL 540
 DB 481 RAISEDOLAPCQWQEHFAERVSFAAPSTDSQGTFRARDQPCATSGPLCLGHL 540
 QY 541 ESSQPPPGF 549
 DB 541 ESSQPPPGF 549

RESULT 5
 DB6_HUMAN STANDARD; PRT; 1406 AA.
 ID DB6_HUMAN STANDARD; PRT; 1406 AA.
 AC P35125; Q15634; O86WP6; O81WT4;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Ubiquitin carboxyl-terminal hydrolase 6 (EC 3.1.2.15) (Ubiquitin
 DE thiolesterase 6) (Ubiquitin-specific processing protease 6)
 DE (Deubiquitinating enzyme 6) (Proco-oncogene TRB-2).
 GN Name=USP6; Synonyms=TRB2;

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.
 RP TISSUE=lung sarcoma;
 RX MEDLINE=92228503; PubMed=1565468;
 RA Nakamura T., Hillova J., Mariage-Samson R., Ono M., Huebner K.,
 RA Camizzone L.A., Boshorian-Sell L., Croce C.M., Hill M.,
 RT "A novel transcriptional unit of the tre oncogene widely expressed in
 RT human cancer cells."
 RL Oncogene 7:733-741(1992).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING, TISSUE
 RP SPECIFICITY, AND DISCUSSION OF TRE2 EVOLUTION
 RX MEDLINE=22506395; PubMed=12604796; DOI=10.1073/pnas.0437015100;
 RA Paulding C.A., Ruvolo M., Haber D.A.,
 RT "The Tre2 (USP6) oncogene is a hominoid-specific gene."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:2507-2511(2003).
 RN [3]
 RN CHARACTERIZATION.
 RP MEDLINE=94067315; PubMed=8247125;
 RA Papa F.R., Hochreiter M.,
 RT "The Yeast DOA4 gene encodes a deubiquitinating enzyme related to a
 RT product of the human tre-2 oncogene."
 RL Nature 366:313-319(1993).
 RN [4]
 RN MUTAGENESIS OF THR-150 AND ARG-187.
 RP PubMed=14521938;
 RA Bitumngu C., De Neve N., Burny A., Bach S., Bontemps F.,
 RA Portetelle D., Vandenbol M.,
 RT "Expression in a RabGAP yeast mutant of two human homologues, one of
 RT which is an oncogene."
 CC Biochem. Biophys. Res. Commun. 310:498-504(2003).
 CC -1- FUNCTION: Has an ATP-independent isopeptidase activity, cleaving
 CC but not isoform 3, shows deubiquitinating activity.
 CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
 CC ubiquitin + a thiol.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=P35125-1; Sequence=Displayed;
 CC Name=2; Synonyms=213(ORF2);
 CC IsoId=P35125-2; Sequence=VSP_010878, VSP_010879;
 CC Name=3; Synonyms=210(ORF1), onctre210p;
 CC IsoId=P35125-3; Sequence=VSP_010880, VSP_010881;
 CC Note=Was shown to be tumorigenic in transfected mice and seems
 CC not to act as GTPase activating protein;
 CC -1- TISSUE SPECIFICITY: Testis specific. Expressed in various cancer
 CC cell lines.
 CC -1- DISEASE: Involved in tumorigenesis.
 CC -1- MISCELLANEOUS: The USP6 gene only exists in the primate lineage.
 CC -1- SIMILARITY: Belongs to peptidase family C19.
 CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; X63546; CAA45108.1; -.
 DR EMBL; X63547; CAA45111.1; -.
 DR EMBL; AY143550; AAN38838.1; -.
 DR EMBL; AY163314; AAO21348.1; -.
 DR PIR; S57867; S57867.
 DR PIR; S57868; S22158.
 DR PIR; S57874; S22155.
 DR MEROPS; C19.009; -.


```
DR Genew; HGNC:12629; USP6.
DR MIM; 604334; -.
DR GO; GO:0005764; C-lysosome; TAS.
DR GO; GO:0004197; F-cysteine-cyste endopeptidase activity; TAS.
DR GO; GO:0003676; F-nucleic acid binding; TAS.
DR InterPro; IPR001394; Peptidase_C19.
DR InterPro; IPR000195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR Pfam; PF00443; UCH; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PSS0086; TBC_RABGAP; 1.
DR PROSITE; PSS0972; UCH_2_1; 1.
DR PROSITE; PSS0973; UCH_2_2; 1.
DR PROSITE; PSS0235; UCH_2_3; 1.
DR Alternative splicing; Hydrolase; Multigene family; Proto-oncogene;
KM Thiol protease; Ubl conjugation pathway.
FT DOMAIN 100 292
FT ACT_SITE 755 755 Rab-GAP TBC.
FT ACT_SITE 1320 1320 By similarity.
FT ACT_SITE 1328 1328 By similarity.
FT VARSPLIC 1 317 Missing (in isoform 2).
FT VARSPLIC 318 359 /FTId=VSP_010878.
FT FT MARLRNQFPTWAMNDVTAKHLRASTKLTTRKQGLDLP
FT FT P -> MPQRLPHARQHTPLGSDYRRVVSVPQPHNDP
FT FT KDSRDA (in isoform 2).
FT FT /FTId=VSP_010879.
FT FT NRPQDNQKQLSV -> ISPLHLQMECSP (in
FT FT isoform 3).
FT FT /FTId=VSP_010880.
FT FT VARSPLIC 787 1406 Missing (in isoform 3).
FT FT /FTId=VSP_010881.
FT FT MUTAGEN 150 150 T->R: Does not restore GAP activity in
FT FT MUTAGEN 187 187 yeast complementation assay.
FT FT R->Q: Does not restore GAP activity in
FT FT CONFLICT 475 475 W -> R (in Ref. 1).
FT FT CONFLICT 912 912 R -> Q (in Ref. 1).
FT FT CONFLICT 963 963 N -> I (in Ref. 1).
SQ SEQUENCE 1406 AA; 158657 MW; D3A6822CEB441DB3 CRC64;

Query Match 73.4%; Score 2188.5; DB 1; Length 1406;
Best Local Similarity 81.6%; Pred. No. 4.3e-129;
Matches 408; Conservative 32; Mismatches 59; Indels 1; Gaps 1;

QY 1 MDVVEAGSWMAQBERDIIMKTEKGRAGLPEDKGPFRSYNNVNDHIGIVHETLPL 60
DB 1 MDVVENADSLQAEKRDILMKYDKGRAGLPEDKGPFRV -GINSSIDRGILHETLPLPV 59
QY 1 TAREAKQIRREISRSKSWVDMLGMEKYSRKLIDRAYKGMPMNIRGPMVSVLNIIEEM 120
DB 60 TAREAKKIRREMTTRTSKMMEMLGEMETYGHSKLIIDRVYKGIIPNIRGPMVSVLNIIOEI 119
QY 121 KLNKNGRVOIMEKGRSSEHIQRIDRVSGTLRKHIIPRDYGTQKORELHILAYEY 180
DB 120 KLNKNGRVOIMEKGRSSEHIQHIDRVTLRNHVFRDRYGAKQRELFTILAYSEY 179
QY 181 NPEVGYCRDLSHIALFLYLIPBEDAFMALVOLLASERHSLOGFSPNGCTVQGLDDOE 240
DB 180 NPEVGYCRDLSHITLFLYLIPBEDAFMALVOLLASERHSLOGFSPNGCTVQGLDDOE 239
QY 241 HVAATSOPKTMGHODKDLCCGCSPLGCLIRLIDGISGLTLRLMDVYLVEGEOLMFI 300
DB 240 HVAAPSOPKTMHODKEGLCCGCSASLGCLIRNLIDGISGLTLRLMDVYLVEGEOLMFI 299
QY 301 TRIARVQOKRLTKTSRCGPWAFRCNRPVDTWARDEDTVLKHLRASMKKLTRKQGLDLP 360
DB 300 TSIALKVOQKRLMKTSRCGLMARLRNQFPDTWAMNDVTVLKHLRASTKLTTRKQGLDLP 359
QY 361 AKPEGSSASRPVPSRSGKTKLCKGDRQAPPGPPAPFPPIPSASPPRAPRSTPCPGA 420
DB 360 ARREGOSLAPRPVPSRSGKTKLCKGYRQAPPGPPAPFPPIPSASPPRAPRSTPCPGA 419
QY 421 VREDTYPVGTGCVPSALAQGGPQSGSWRLQWNSWMLPTDIDVEGPFMRHYDFRQSCWV 480
```

```
DB 420 VREDTYPVGTGCVPSALAQGGPQSGSWRLQWNSWMLPTDIDVEGPFMRHYDFRQSCWV 479
QY 481 RAISOEDOLAPCWOAEHPAE 500
DB 480 RAISOEDOLATCWOAEHCE 499

RESULT 6
QI5635
ID QI5635 PRELIMINARY; PRT; 376 AA.
AC QI5635;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLREL. 24, Last annotation update)
DE Oncogene.
GN Name=tre;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Esing's sarcoma;
RX MEDLINE=92228503; PubMed=1565468;
RA Nakamura T., Hilla J., Marriage-Samson R., Ono M., Huebner K.,
RA Cantizaro L.A., Boghosian-Sell L., Croce C.M., Hill M.;
RT "A novel transcriptional unit of the tre oncogene widely expressed in
human cancer cells."
RL Oncogene 7:733-741(1992).
DR EMBL; X63547; CAA45110.1; -.
DR PIR; S57867; S57867.
DR InterPro; IPR000195; RabGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PSS0086; TBC_RABGAP; 1.
SQ SEQUENCE 376 AA; 43160 MW; F43A86D09F344117 CRC64;

Query Match 51.8%; Score 1544.5; DB 2; Length 376;
Best Local Similarity 77.9%; Pred. No. 3.1e-89;
Matches 293; Conservative 29; Mismatches 51; Indels 3; Gaps 2;

QY 1 MDVVEAGSWMAQBERDIIMKTEKGRAGLPEDKGPFRSYNNVNDHIGIVHETLPL 60
DB 1 MDVVENADSLQAEKRDILMKYDKGRAGLPEDKGPFRV -GINSSIDRGILHETLPLPV 59
QY 61 TAREAKQIRREISRSKSWVDMLGMEKYSRKLIDRAYKGMPMNIRGPMVSVLNIIEEM 120
DB 60 TAREAKKIRREMTTRTSKMMEMLGEMETYGHSKLIIDRVYKGIIPNIRGPMVSVLNIIOEI 119
QY 121 KLNKNGRVOIMEKGRSSEHIQRIDRVSGTLRKHIIPRDYGTQKORELHILAYEY 180
DB 120 KLNKNGRVOIMEKGRSSEHIQHIDRVTLRNHVFRDRYGAKQRELFTILAYSEY 179
QY 181 NPEVGYCRDLSHIALFLYLIPBEDAFMALVOLLASERHSLOGFSPNGCTVQGLDDOE 240
DB 180 NPEVGYCRDLSHITLFLYLIPBEDAFMALVOLLASERHSLOGFSPNGCTVQGLDDOE 239
QY 241 HVAATSOPKTMGHODKDLCCGCSPLGCLIRLIDGISGLTLRLMDVYLVEGEOLMFI 300
DB 240 HVAAPSOPKTMHODKEGLCCGCSASLGCLIRNLIDGISGLTLRLMDVYLVEGEOLMFI 299
QY 301 TRIARVQOKRLTKTSRCGPWAFRCNRPVDTWARDEDTVLKHLRASMKKLTRKQGLDLP 360
DB 300 TSIALKVOQKRLMKTSRCGLMARLRNQFPDTWAMNDVTVLKHLRASTKLTTRKQGLDLP 359
QY 361 AKPEGSS--ASRPV 374
DB 360 GPTALGRCVAGSP 375

RESULT 7
Q6PD72
```

ID AC Q6BD72 PRELIMINARY; PRT; 346 AA.
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Tothiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton A., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smaltus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058890; AAH58890.1; -
 DR InterPro; IPR000195; RABGAP_TBC.
 DR Pfam; PF00566; TBC; 1.
 DR SMART; SM00164; TBC; 1.
 DR PROSITE; PS50086; TBC_RABGAP; 1.
 DR KW Hypothetical protein.
 SQ SEQUENCE 346 AA; 39686 MW; ADBF92563358FCBF CRC64;
 Query Match
 Best Local Similarity 49.6%; Score 1479; DB 2; Length 346;
 Matches 276; Conservative 13; Mismatches 13; Indels 0; Gaps 0;
 QY 1 MDVVEVAGSWMAOEREDIMKYEKGRAGLPEDKGPFPFSYNNVNDHGIYHETELPPL 60
 DB 1 MDVVEVAGSWMAOEREDIMKYEKGRAGLPEDKGPFPFSYNNVNDHGIYHETELPPL 60
 QY 61 TAREAKOIRREISRSKSWVDMLGDWEKYSRRKLDRAVKGPMNIRGPMWSVLNTEEM 120
 DB 61 TAREAKOIRREISRSKSWVDMLGDWEKYSRRKLDRAVKGPMNIRGPMWSVLNTEEM 120
 QY 121 KLNKPGRYOIMKEKGRSSSEHIQRIDRVSGTLRHGIFFRDRYGTQKRELLHLLAYEY 180
 DB 121 KLNKPGRYOIMKEKGRSSSEHIQRIDRVSGTLRHGIFFRDRYGTQKRELLHLLAYEY 180
 QY 181 NPEVGYCRDLISHAALFLYLPEEDAFWALVOLASERHSLQGFHSPNGTVOGLODOOE 240
 DB 181 NPEVGYCRDLISHAALFLYLPEEDAFWALVOLASERHSLQGFHSPNGTVOGLODOOE 240
 QY 241 HVVATSOPTMGHODKOLCGQCSPLGCLIRLIDISIGLTLRLMDVYLVEGQALMPI 300
 DB 241 HVVATSOPTMGHODKOLCGQCSPLGCLIRLIDISIGLTLRLMDVYLVEGQALMPI 300
 QY 301 TR 302
 DB 301 TR 302

DB 301 TR 302
 RESULT 8
 ID AAH58890 PRELIMINARY; PRT; 346 AA.
 AC AAH58890;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Tothiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton A., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Krzywinski M.I., Skalska U., Smaltus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC058890; AAH58890.1; -
 DR KW Hypothetical protein.
 SQ SEQUENCE 346 AA; 39686 MW; ADBF92563358FCBF CRC64;
 Query Match
 Best Local Similarity 49.6%; Score 1479; DB 2; Length 346;
 Matches 276; Conservative 13; Mismatches 13; Indels 0; Gaps 0;
 QY 1 MDVVEVAGSWMAOEREDIMKYEKGRAGLPEDKGPFPFSYNNVNDHGIYHETELPPL 60
 DB 1 MDVVEVAGSWMAOEREDIMKYEKGRAGLPEDKGPFPFSYNNVNDHGIYHETELPPL 60
 QY 61 TAREAKOIRREISRSKSWVDMLGDWEKYSRRKLDRAVKGPMNIRGPMWSVLNTEEM 120
 DB 61 TAREAKOIRREISRSKSWVDMLGDWEKYSRRKLDRAVKGPMNIRGPMWSVLNTEEM 120
 QY 121 KLNKPGRYOIMKEKGRSSSEHIQRIDRVSGTLRHGIFFRDRYGTQKRELLHLLAYEY 180
 DB 121 KLNKPGRYOIMKEKGRSSSEHIQRIDRVSGTLRHGIFFRDRYGTQKRELLHLLAYEY 180
 QY 181 NPEVGYCRDLISHAALFLYLPEEDAFWALVOLASERHSLQGFHSPNGTVOGLODOOE 240
 DB 181 NPEVGYCRDLISHAALFLYLPEEDAFWALVOLASERHSLQGFHSPNGTVOGLODOOE 240
 QY 241 HVVATSOPTMGHODKOLCGQCSPLGCLIRLIDISIGLTLRLMDVYLVEGQALMPI 300
 DB 241 HVVATSOPTMGHODKOLCGQCSPLGCLIRLIDISIGLTLRLMDVYLVEGQALMPI 300
 QY 301 TR 302
 DB 301 TR 302

Db 301 TK 302

RESULT 9

06PIJ2 PRELIMINARY; PRT; 376 AA.

AC 06PIJ2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DE Similar to TBC1 domain family, member 3.
GN Name=MGCC44903;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033670; AAH33670.1;
DR InterPro; IPR000195; RABGAP_TBC.
DR Pfam; PF00566; TBC; 1.
DR SMART; SM00164; TBC; 1.
DR PROSITE; PS50086; TBC_RABGAP; 1.
SQ SEQUENCE 376 AA; 42391 MW; F3A9AC94959B9900 CRC64;

Query Match 49.5%; Score 1477.5; DB 2; Length 376;
Best Local Similarity 72.7%; Pred. No. 5.1e-85;
Matches 301; Conservative 12; Mismatches 40; Indels 61; Gaps 8;

QY 1 MDVVEASGWAQOEEDIIIMKYEKGHRAGLPEDKGPFPFRSYNNNVNDHIGIVHETELPPL 60
DB 1 MDVVEASGWAQOEEDIIIMKYEKGHRAGLPEDKGPFPFRSYNNNVNDHIGIVHETELPPL 60
QY 61 TAREKQJRIETSRKSKVNDMLGDWEKYSRKLIDRAYKGMPPNIRGPMVSLNIEEM 120
DB 61 TAREKQJRIETSRKSKVNDMLGDWEKYSRKLIDRAYKGMPPNIRGPMVSLNIEEM 120
QY 121 KLNKNGRYOIMKEKRSSEHIQRIIDRVSGTLARKHIFPRDRYGTGKRELLHILLAYEY 180
DB 121 KLNKNGRYOIMKEKRSSEHIQRIIDRVSGTLARKHIFPRDRYGTGKRELLHILLAYEY 180
QY 161 NPEVGYCDLSHIALFLIYLPEEDAFAVALVOLLASERHSLQGFSPNGCTVQGLDDOGE 240
DB 161 NPEVGYCDLSHIALFLIYLPEEDAFAVALVOLLASERHSLQGFSPNGCTVQGLDDOGE 240

QY 241 HVATSQPTMGHODKDCGQSPGLIRLIDGISLGLTRLMVDVYVEGE-----QA 296
DB 241 HVATSQPTMGHODKDCGQSPGLIRLIDGISLGLTRLMVDVYVEGE-----QA 296
QY 297 LMP-----ITRIAFKVGQKRLTKTSRCGPWAFPCNRFVDTWARD 337
DB 297 LMP-----ITRIAFKVGQKRLTKTSRCGPWAFPCNRFVDTWARD 337
QY 338 TVLKLRLASMKLTKQGDLPPEAKPEQSSASRPVPSRGKTKLCKGRDAP 391
DB 338 TVLKLRLASMKLTKQGDLPPEAKPEQSSASRPVPSRGKTKLCKGRDAP 391
QY 347 -----TLERAQSLKPMASP-----SLQEPHPS-----NECP 376
DB 347 -----TLERAQSLKPMASP-----SLQEPHPS-----NECP 376

RESULT 10

AAH33670 PRELIMINARY; PRT; 376 AA.

AC AAH33670;
DT 01-JUN-2004 (TRENBLrel. 27, Created)
DT 01-JUN-2004 (TRENBLrel. 27, Last sequence update)
DE TBC1D3 protein.
GN TBC1D3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033670; AAH33670.1;
SQ SEQUENCE 376 AA; 42391 MW; F3A9AC94959B9900 CRC64;

Query Match 49.5%; Score 1477.5; DB 2; Length 376;
Best Local Similarity 72.7%; Pred. No. 5.1e-85;
Matches 301; Conservative 12; Mismatches 40; Indels 61; Gaps 8;

QY 1 MDVVEASGWAQOEEDIIIMKYEKGHRAGLPEDKGPFPFRSYNNNVNDHIGIVHETELPPL 60
DB 1 MDVVEASGWAQOEEDIIIMKYEKGHRAGLPEDKGPFPFRSYNNNVNDHIGIVHETELPPL 60
QY 61 TAREKQJRIETSRKSKVNDMLGDWEKYSRKLIDRAYKGMPPNIRGPMVSLNIEEM 120
DB 61 TAREKQJRIETSRKSKVNDMLGDWEKYSRKLIDRAYKGMPPNIRGPMVSLNIEEM 120
QY 121 KLNKNGRYOIMKEKRSSEHIQRIIDRVSGTLARKHIFPRDRYGTGKRELLHILLAYEY 180
DB 121 KLNKNGRYOIMKEKRSSEHIQRIIDRVSGTLARKHIFPRDRYGTGKRELLHILLAYEY 180
QY 161 NPEVGYCDLSHIALFLIYLPEEDAFAVALVOLLASERHSLQGFSPNGCTVQGLDDOGE 240
DB 161 NPEVGYCDLSHIALFLIYLPEEDAFAVALVOLLASERHSLQGFSPNGCTVQGLDDOGE 240

86Dcb4		
ID		
AC	OedCB4	PRELIMINARY;
OC	OedCB4;	PRT; 291 AA.
DT	01-OCT-2004 (TREMBLrel. 28,	
BT	01-OCT-2004 (TREMBLrel. 28,	Created)
DT	01-OCT-2004 (TREMBLrel. 28,	Last sequence update)
DE	Hypothetical protein.	
OS	Homo sapiens (Human).	
CC	Eukaryota; Metazoa; Chordata;	
OC	Mammalia; Eutheria; Primates;	
NBFI	NCBI_TaxId=9606;	Catarrhini, Hominiidae, Homo.
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Braint	

PubMed=12477932;
Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bat N.K.,
Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hasle P.,
Datchenko L., Marusina K., Farmer A.C., Rubin G.M., Hong L.,
Sapichenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
Brownstein M.J., Usdin T.B., Rosilyuk S., Carninci P., Prange C.,
Rahm S.A., Loquellano N.A., Peters G.U., Abramson R.D., Mullaly S.J.,
Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
Valliano D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs K.A.,
Falley U., Helton E., Ketteman M., Maden A.C., Rodrigues S., Sanchez A.,
Whiting R., Maden A., Young A.C., Maden A.C., Rodriguez S., Sanchez A.,
Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywniski M.I., Skalska U., Smalins D.E., Schermer A., Schein J.E.,
Jones S.J., Matra M.A.;
Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences." ;
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Strandberg R.;
Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
EMBL; BC078140; AAH78140.1; -.
Hypothetical protein.
SEQUENCE 201 nt.

Query Match	Score	DB 2	Length
Test Local Similarity	45.1%	1346	291
Matches 252; Conservative	99.2%	Pred. No. 7.2e-77	
		1; Mismatches	

1 MDVVEVAGSMAQEREDIMKKEKGRAGLPEDGKPPFRSYNNVNDHGIYHETELPPL 60
61 TAREAKOTPPPTCRGCTT-----

120

	RESULT	12
D	US6L_MOUSE	
C	US6L_MOUSE	
X	080XC3;	STANDARD; PRT; 819 AA.
T	01-OCT-2004 (Rel. 45, Created)	
T	01-OCT-2004 (Rel. 45, Last sequence update)	
T	01-OCT-2004 (Rel. 45, Last annotation update)	
N	Name=Usp6n1; Synonyms=Klaa0019;	
S	Mus musculus (Mouse).	
C	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
C	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;	
X	NCBI_TaxID=10090;	
[1]		

TISSUE=Embryonic tail;
 PubMed=14621295;
 Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 Suga Y., Nagase T., Ohara O., Koga H.,
 "Prediction of the coding sequences of mouse homologues of XIRA gene:
 III The complete nucleotide sequences of 500 mouse XIRA-homologues
 cDNAs identified by screening of terminal sequences of cDNA clones
 randomly sampled from size-fractionated libraries.",
 DNA Res. 10:167-180 (2003).
 (2)
 SEQUENCE FROM N.A. (ISOFORM 2).
 TISSUE=Eye;
 MEDLINE=22388257; PubMed=12477937; DOI=10.1073/pnas.242603899;
 Krausberg R.L., Feingold R.A., Grouse L.H., Derge J.,
 Altschul S.F., Zeeberg B., Wagner L., Shemem C.M., Schuler G.D.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 DiCicco L., Marnett K., Farmer A., Rubin G.W., Hsieh F.,
 Stappleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 Rata S.S., Louie T.B., Toshiyuki S., Carninci P., Prange C.,
 Bosak S.A., Melnan P.U., McKernan K.J., Abramson R.D., Millaby S.J.,
 Richards S.,

Phaley J., Helton E., Kettner, M.M., Sodergren E.J., Lu X., Gibbs R.A., Whiting J., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitt J., Wiers D.M., Rutherford J. v S. N.

11- FOUNCTION: Actress
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." ,
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002)
11- FOUNCTION: Actress
"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences." ,
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002)

- 1- SUBUNIT: Interacts with EPS8 (By similarity).
- 1- ALTERNATIVE PRODUCTS: Event=Alternative splicing.

```
Isold=Q80XC3-1; SequenceDisplayed,  
Name=2;  
Isold=Q80XC3-2; Sequence-ver
```

```

CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 671.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AK129037; BAC97847.1; ALT INIT.
CC EMBL: BC051184; AAH51184.1; ALT_FRAME.
CC MGI: 2138693; A1316785.
CC InterPro: IPR000195; RabGAP_TBC.
CC DR Pfam: PF00566; TBC; 1.
CC DR SMART: SM00164; TBC; 1.
CC DR PROSITE: PS00086; TBC_RABGAP; 1.
CC KW Alternative splicing; GTPase activation.
CC FT DOMAIN 100 292 Rab-GAP TBC.
CC FT VARSPLIC 1 2 Pro-rich.
CC FT NM -> MIQVQLVLELTPSRQKATAKED (in
CC FT isoform 2).
CC FT /FTID=VSP_011153.
CC SQ SEQUENCE 819 AA; 93573 MW; AD6B72BE42A00B54 CRC64;

Query Match 22.0%; Score 655; DB 1; Length 819;
Best Local Similarity 34.4%; Pred. No. 7.4e-33;
Matches 156; Conservative 72; Mismatches 191; Indels 34; Gaps 9;

OY 12 AGEREDIMKYEKGKRAGLPEDKGPFRPSYNNVNDHIGIVHETLPLTAEAKQIRRE 71
DB 12 AGEREIVAKVDGRGEGAEIEPWEDADYLVY-KVTDREGFLHEBELPYHNAADROKOLE 70
OY 72 ISRSKQVMDLDMKYSRKLIDRAYKGMNIRGPMWSVLTNIEMKLNKPGRYQIM 131
DB 71 IERTSKMLKMKWRYKNTKFKHRIYKGIPLQIRGEVWALLLEIPKKESTRDLYSKL 130
OY 132 KEKGRSEHRIORIDVSGTLRKHIIPRDRTGTQRELLHLLAYEEYNPEVGYCRDLS 191
DB 131 KHRAGCSPDIRQIDLDVNRTRFDHIMFRDRYGVQOQLFHYLAAYSYNTEVGYCOGMS 190
OY 132 KEKGRSEHRIORIDVSGTLRKHIIPRDRTGTQRELLHLLAYEEYNPEVGYCRDLS 191
DB 131 KHRAGCSPDIRQIDLDVNRTRFDHIMFRDRYGVQOQLFHYLAAYSYNTEVGYCOGMS 190
OY 192 HIALFLYLPEDEAFMLVOLLASERHSLOGFHSNPGSTVQG-----LDQOEHVAT 245
DB 191 QITALLMYMEDEAFMLVLKLFSGPKHAMGFF-----VQSPFKLRFQHHKELINK 244
OY 246 SQPKTMGHQDKDLCGQCSPLGCLIRILIDGSLGLTLRLMDVYLVEGEOALMPITRIAF 305
DB 245 FLSTKQKHLDSQEIYTSFYTMKMFQCPFLDRTPFLNLRIMWIYIFEGERVLTAMSYTL 304
OY 306 KVOQKRLTKTSRCGWFARFCNRFVDTWARD---EDTVLKHLRASMKKLTKRQGLPPPA 361
DB 305 KLHKHKLMLKLS---MEELVEFLQETLAKDPFEEDFVIEQLQVSMALKRAKLDLPBFG 360
OY 362 KPEQSSASRPV---PASRGKTLCKGDRQ-APPGPAPR-----PPPIWSASPPRAP 410
DB 361 KEDE--YPKKPLGQLPPESACVNHLSNGORSVGRPSPTSSRREDGSPRKNHSPVHHS 418
OY 411 RSTPCPGAVREDTPYVGTQGVSPPALAQGP 443
DB 419 RNGTPERAGQSRKRSVDEGSKNLKHEASQRKP 451

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=embryonic tail;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIA Gene:
RT I11. The Complete Nucleotide Sequences of 500 Mouse KIAA-homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-fractionated Libraries."
RL DNA Res. 10:167-180(2003).
DR EMBL: AK129037; BAC97847.1; -.
FT NON_TER 1 1
SQ SEQUENCE 841 AA; 96174 MW; B96ABE416E245C8A CRC64;

Query Match 22.0%; Score 655; DB 2; Length 841;
Best Local Similarity 34.4%; Pred. No. 7.6e-33;
Matches 156; Conservative 72; Mismatches 191; Indels 34; Gaps 9;

OY 12 AGEREDIMKYEKGKRAGLPEDKGPFRPSYNNVNDHIGIVHETLPLTAEAKQIRRE 71
DB 34 AGEREIVAKVDGRGEGAEIEPWEDADYLVY-KVTDREGFLHEBELPYHNAADROKOLE 92
OY 72 ISRSKQVMDLDMKYSRKLIDRAYKGMNIRGPMWSVLTNIEMKLNKPGRYQIM 131
DB 93 IERTSKMLKMKWRYKNTKFKHRIYKGIPLQIRGEVWALLLEIPKKESTRDLYSKL 152
OY 132 KEKGRSEHRIORIDVSGTLRKHIIPRDRTGTQRELLHLLAYEEYNPEVGYCRDLS 191
DB 153 KHRAGCSPDIRQIDLDVNRTRFDHIMFRDRYGVQOQLFHYLAAYSYNTEVGYCOGMS 212
OY 192 HIALFLYLPEDEAFMLVOLLASERHSLOGFHSNPGSTVQG-----LDQOEHVAT 245
DB 213 QITALLMYMEDEAFMLVLKLFSGPKHAMGFF-----VQSPFKLRFQHHKELINK 266
OY 246 SQPKTMGHQDKDLCGQCSPLGCLIRILIDGSLGLTLRLMDVYLVEGEOALMPITRIAF 305
DB 267 FLSTKQKHLDSQEIYTSFYTMKMFQCPFLDRTPFLNLRIMWIYIFEGERVLTAMSYTL 326
OY 306 KVOQKRLTKTSRCGWFARFCNRFVDTWARD---EDTVLKHLRASMKKLTKRQGLPPPA 361
DB 327 KLHKHKLMLKLS---MEELVEFLQETLAKDPFEEDFVIEQLQVSMALKRAKLDLPBFG 382
OY 362 KPEQSSASRPV---PASRGKTLCKGDRQ-APPGPAPR-----PPPIWSASPPRAP 410
DB 363 KEDE--YPKKPLGQLPPESACVNHLSNGORSVGRPSPTSSRREDGSPRKNHSPVHHS 440
OY 411 RSTPCPGAVREDTPYVGTQGVSPPALAQGP 443
DB 441 RNGTPERAGQSRKRSVDEGSKNLKHEASQRKP 473

```

```

RESULT 14
US6L_HUMAN STANDARD; PRT; 828 AA.
ID US6L_HUMAN Q92738; Q15400;
AC Q92738; Q15400;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE US6 N-terminal like protein (Related to the N terminus of tre) (RN-
DE cre).
GN Name=USP6NL; Synonyms=KIAA0019;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND INTERACTION WITH BP98.
RX MEDLINE=96293402; PubMed=8700515;
RA Wong W.T., Seki N., Nagase T., Nomura N., Robbins K.C., Di Fiore P.P.,

```

RA Matoskova B.;
 RT "RN-tre identifies a family of tre-related proteins displaying a novel
 RL potential protein binding domain.",
 RN Oncogene 12:2563-2571(1996).
 RP [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX TISSUE=Testis;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shellen G.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stedman M., Soares M.B., Donaldson M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mallaby S.J.,
 RA Rich S.A., McGowan P.J., McKernan K.J., Garavito R.M., Hong L.,
 RA Villalon D.K., Murty K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield V.S.N., Krzyzinski M.I., Skalska U., Smalins D.E.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP FUNCTION, AND MUTAGENESIS OF ARG-106; ASP-147 AND ARG-150.
 RA PubMed=11099046; DOI=10.1038/35042605;
 RA Lanzetta L., Rybin V., Malabarida M.G., Christoforidis S., Scita G.,
 RA Zerial M., Di Fiore P.P.;
 RT "The EphA protein coordinates EGF receptor signalling through Rac and
 RL Nature 408:374-377(2000).
 CC -1- FUNCTION: Acts as a GTPase activating protein for RAB5A. Involved
 CC in receptor trafficking. In complex with EPS8 inhibits
 CC internalization of EGFR.
 CC -1- SUBUNIT: Interacts with EPS8.
 CC -1- TISSUE SPECIFICITY: Widely expressed.
 CC -1- SIMILARITY: Contains 1 Rab-GAP TBC domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 DR EMBL; D13644; BAA02807.2; ALT_INT.
 DR EMBL; BC042943; AAH42943.1; --
 DR Genbank; HGNC:16858; USP6M.
 DR MIM; 605405; --
 DR InterPro; IPR000195; RabGAP_TBC.
 DR Pfam; PF00566; TBC; 1.
 DR SMART; SM00164; TBC; 1.
 DR PROSITE; PS50086; TBC_RABGAP; 1.
 KW GTPase activation.
 FT DOMAIN 100 292 Rab-GAP TBC.
 FT MUTAGEN 106 106 R->A: Loss of GAP activity on RAB5A.

FT MUTAGEN 147 147 D->A: Loss of GAP activity on RAB5A.
 FT MUTAGEN 150 150 R->A: Loss of GAP activity on RAB5A.
 SQ SEQUENCE 828 AA, 94104 MW, 45E3DBDFEE04E78 CRC64;
 Query Match
 Best Local Similarity 33.6%; Pred. No. 1,36-32;
 Matches 164; Conservative 72; Mismatches 172; Indels 80; Gaps 14;
 QY 12 AGEREDIMKKEKGRAGLDEDPGPPFRSYNNNDVHLGIYETELPPLTAREAKQIRRE 71
 DB 12 AGEREATVAYKXDRGREGALTEPWEADYLVY-KVDRFGFLHEBELDHNVAVERQKMLE 70
 QY 72 ISRSKXVMDLGPWEKXSSRLIDRAYKGMNIRGPMNVLLNTEMLKNPGRYOIM 131
 DB 71 IERTTKMLKMLKWEKXKXNTEKFRIRYKGIPLDRGEVWMLLEIPKMEETRDLYSKL 130
 QY 132 KKKKSSSEHIGRIDRVSGTLRKHIIPRDYGTQKQELHILLAYEYVPEVGYCRDLS 191
 DB 131 KHRARCSFDIRQIDLVNRTFRDHIMFRDYGKQOOLPHVLAAYSYNTEVGYCQCMS 190
 QY 192 HIALFLYLPREDAPFALVOLLASERHSLQGFHSPNGVVOG-----LDQOEHVAT 245
 DB 191 QTTALLVYVMEEDHFMALVLFSGPKGNHGF-----VGFPLKLRPFQEHHEKTLNK 244
 QY 246 SQPKMGQDKDLCQGCSPGLGLIRILIDGSLGLTLMNVLYVEEQLMPTIRIAF 305
 DB 245 FLSTKKQHLDSQELIYTSFYTMKWFPGCFIDRTPFLNLRIMDIYIFEGERVLTAMSYTIL 304
 QY 306 KVQGRRLTKTSRCGPMARFCNRPVDWARD-----EDYVLRKLRSMKMLTRKQGDLPSPA 361
 DB 305 KLKHGHLMKLS--MEELVEFQETLADFFEDDFYIETLOISMTLBRATLDIPEP 360
 QY 362 K-----PQGS-----SASRPVPSRSGKTLCKGDAQP----- 390
 DB 361 KEDEYRKPELGQPEPLDSQWGVHLSNGQSRVGRPSPLAG-----RRSGAPRRHENS 415
 QY 391 PGPPAPFPPIVSAAPRPSSTPCGAVRHTYTVVQG-----VSPPALAQSPQG 445
 DB 416 FHKQSTGTP-ERRQPPRRK-----SVESKSLKLEADFPKPLPS-----GQD 459
 QY 446 SRRFLQWN 453
 DB 460 SSR-QYN 465
 RESULT 15
 BAA02807 PRELIMINARY; PRT; 838 AA.
 AC BAA02807;
 DT 02-MAR-2004 (TrEMBLrel. 27, Created)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
 DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
 DE K1A0019 protein (Fragment).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RP MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=6051387; PubMed=7584026;
 RA Nomura N., Miyajima N., Sazuka T., Tanaka A., Kawarabayashi Y.,
 RT "Prediction of the coding sequences of unidentified human genes. I.
 RL The coding sequences of 40 new genes (K1A0001-K1A0040) deduced by
 RN cell line KG-1".
 RN DNA Res. 1:27-35(1994).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=60513

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 05:11:13 ; Search time 6811 Seconds

(without alignments)
3811.782 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVVEVAGSWAQAEREDII.....TSGPCLCGHLHESQGPFGF 549

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=x1h

-C=/cgn2.1/USPTO.epoc/US10071833/runac_03022005_071218_187/app_query.fasta.1.711

-DB=Genembl -QFMT=fastap -SUFFIX=xrge -MINMATCH=0.1 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-POCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=PCT -NORM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10071833@CGN 1_1_3852@runac_03022005_071218_187 -NCPU=6 -ICPU=3

-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genembl:*

1: gb_da:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pal:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sfr:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	6	AX775945 Sequence
2	2983	100.0	1964	6	AF540953 Homo sapi
3	2975	99.7	1964	6	AX775947 Sequence
4	2971	99.6	2090	9	BC071680 Homo sapi

5	2963	99.3	2072	6	AX086847 Sequence
6	2963	99.3	2072	9	HSM801828
7	2945	98.7	2166	9	BC075809 Homo sapi
8	2906.5	97.4	2304	6	AX775943 Sequence
9	2898.5	97.2	2304	6	AX775941 Sequence
10	2849	95.5	2575	6	BC033670 Homo sapi
11	2835	95.0	2647	6	AX775951 Sequence
12	2827	94.8	2647	6	AX775949 Sequence
13	2827	94.8	2647	6	AX833552 Sequence
14	2827	94.8	2647	6	AX095385 Homo sapi
15	2732	91.6	2111	9	BC058890 Homo sapi
16	2626.5	88.0	1924	9	BC078140 Homo sapi
17	2595	87.0	3389	9	AX122833 Homo sapi
18	2191.5	73.5	7878	6	AX332512 Sequence
19	2191.5	73.5	7878	6	HSRRE210
20	2188.5	73.4	7971	9	AY143550 Homo sapi
21	2184.5	73.2	8491	9	HSW807865
22	2133.5	71.5	8201	6	I76205
23	2133.5	71.5	8201	6	AX774956
24	2133.5	71.5	8201	9	HSRRE213
25	1532	51.0	134506	2	AC146482 Homo sapi
26	1532	51.0	180876	9	AC021317
27	1532	51.0	181517	9	AC126327
28	1521	51.0	193646	9	AC131056 Homo sapi
29	1518	50.9	146243	2	AC130293
30	1514	50.8	37757	2	AC139091 Homo sapi
31	1505	50.5	173053	2	AC129893
32	1501	50.3	140152	9	AC027821
33	1500.5	50.3	140152	9	AC027821
34	1500.5	50.3	158698	2	AC036181
35	1500.5	50.3	192881	9	AC067923
36	1500	50.3	206040	2	AC132814
37	1497.5	50.2	185449	2	AC139361
38	1494.5	50.1	42073	2	AC144840
39	1493.5	50.1	41266	2	AC139176
40	1493.5	50.1	89147	2	AC146857 Homo sapi
41	1487.5	49.9	146243	2	AC130293
42	1469	49.2	123925	9	AC003976 Homo sapi
43	1465	49.1	123925	2	AC023133
44	1448.5	48.6	163628	2	AC139367
45	1373.5	46.0	36160	2	AC139089 Homo sapi

ALIGNMENTS

RESULT 1	AX775945	Sequence 215 from Patent WO03048202.	1964 bp	mRNA	linear	PAT 14-JUL-2003
LOCUS	AX775945					
DEFINITION	AX775945					
ACCESSION	AX775945.1	GI:32693663				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
REFERENCE	1	Matsuda, A. and Muramatsu, S.				
AUTHORS		NF-KB activating gene				
TITLE		Patent: WO 03048202-A 215 12-JUN-2003;				
JOURNAL		Asahi Kasei Kabushiki Kaisha (JP)				
FEATURES		Location/Qualifiers				
source		1..1964				
		/organism="Homo sapiens"				
		/mol_type="mRNA"				
		/db_xref="taxon:9606"				
		1..1650				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="CAE11672.1"				
		/db_xref="GI:32693664"				
		/translation="MDVVEVAGSWAQAEREDIIIMKYEGKHAGLPEDKGPFRSYNN				
		NVDHGIYHETELPPLTAREAKQIRREISRSKSWDMDGDEKYSRKLIDRAVYGM				

Rubin, M., Hoey, T., Powers, S., and Li, J.
 PRC17, a novel oncogene encoding a Rab GTPase-activating protein,
 is amplified in prostate cancer
 Cancer Res. (2002) in press
 2 (bases 1 to 1964)
 Li, J.-L., Peng, Y., and Powers, S.
 Direct Submission
 Submitted (23-AUG-2002) Genomics, Tularik, Inc, 266 E Pulaski Road,
 Suite 1, Greenham, NY 11740, USA

FEATURES

source

gene

CDS

```

1..1964
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q11"
1..1964
/gene="PRC17"
1..1650
/gene="PRC17"
/codon_start=1
/product="Rab GTPase-activating protein PRC17"
/protein_id="AA03117.1"
/db_xref="GI:23452665"
/translation="MDVVEVAGSMVAOEREDIMKYEKGRAGLPEDKPKPESYN
NDHLGIHETELPLTLRAKQIRBEISRKSWDMGDMKYSRKLIDRAKGM
PMNIRGPMASVLTINEMKLNKRGVQIKRKGRSSHHIOIDRDSGTLLKHIKPR
DRYGRKORELHITLAEYENPVGICRDLSHIALFLYLBEDAPALVOLLASER
HSLQGFHSPNGSTVGLQDQDQSHVATISQPKTMGHQDKKDLGQCSPLGLRIIDG
ISLGLRLMDVLYVEGGLAMPITRIAFKVKQKRLTKTSRCGPWAFRPNRVDIAR
DEDVTLKLRASMKLTKRQDLPAPAKESGSSASRPVAPSGKTKCKGRQAPG
PAPRPRIKMSPPRAPRPSSTPCGAVREDTYPVGQGVSPALAOGPGGWSRPL
QMSNMPRLPTDIDVEGPMFRHYDPDROSQWVRAISOEDOLAPCQWQEHPRERVSAPAA
PSTDSQGTFFRANDEQPCAPTSFPLCGLHSSQFPFGF"

```

ORIGIN

Alignment Scores:

Pred. No.: 1,126-103 Length: 1964
 Score: 2983.00 Matches: 549
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-10-071-838-2 (1-549) x AF540953 (1-1964)

```

QY 1 MetAEPValaIgluValaIagIySerTPTPAlaGIngluArgluAapIleIleMet 20
DB 1 ATGACGTGTGAGAGGTGCGCGGCAAGTTGTGTGGGCAAGAGCGAGACATCATTTATG 60
QY 21 LysIYTGluVgIyHIArGAlaGlyLeuPProGluAapIyGlyPProLysPProPheArg 40
DB 61 AAATCGAAGAGGACACCGAGCTGGCTGCCAGAGGACAGGGGGCTTAAGCCTTTTCA 120
QY 41 SerTYrAsnAspAspValaAspHisIleuGlyIleValHisGluThrcIuLeuPProLeu 60
DB 121 AGCTTACACACACAGCTCATCATTTTGGGATTTGATCATGACAGGAGCTGCTCTCTG 180
QY 61 ThrAlaArgIuAlaIyGInIleArgArgIuIleSerArgIySerIySerIyTPValaAp 80
DB 181 ACTGCGCGGAGCGGAGCAAAATTCGGCGGAGATCAAGCGAAAGAGCAAGTGGTGAT 240
QY 81 MetLeuGlyAapITPGLuLySerIyLysSerSerArgIyLeuIleAapArgAlaTyLys 100
DB 241 ATGCTGGAGAGACTGGAGAAATACAAAGACAGAAAGCTCATGATGACCGTACAG 300
QY 101 GlyMetProMetAsnIleArgIyPProMetTPSerValIleuIleuAenIleGluGluMet 120
DB 301 GGAATGCCCATTAACATCCGGGGCCGATGTGCTCATCTCTGACATTAGAGAAATG 360
QY 121 LysLeuLyAapnProGlyArgTYrGInIleMetLysGluLyGlyLysArgSerSerGlu 140
DB 361 AAGTGAAGAAACCCCGAGAGATACAGATCATGAAGAGAGGAGGCAAGGTCACTGAG 420

```

```

QY 141 HisIleGInArgIleAapArgAapValSerGlyThrcLeuArgLysHisIlePhePheArg 160
DB 421 CACATCCAGCGCATGACACCGGACCTTAAGCGGACATTAGAGAACATATATCTTCCAG 480
QY 161 AapArgTYrGlyThrcLysGlnArgIuLeuLeuHisIleLeuLeuAlaTyGluGluTy 180
DB 481 GATCATAGAGAAACCAAGACGCGGAGACTACATCCATCTCTGCTCATATAGAGATAT 540
QY 181 AapProGluValaIyTYrCyArGAspIleuSerHisIleAlaIleuPheLeuLeuTy 200
DB 541 AACCCGAGGTGGCTACTGACAGGACCTGAGCCACATCGCGCTTGTCTCTCAT 600
QY 201 LeuPProGluGluAapAlaPheTPAlaLeuValGInLeuLeuAlaSerGluAapIleSer 220
DB 601 CTTCCTGAGAGAGATGATCTTGGGCACTGGTGCAGCTGTGGCCAGTACAGAGCACTCC 660
QY 221 LeuGInGlyPheHisSerProAapnIyGlyThrcValGInGlyLeuGlnAapGInGlu 240
DB 661 CTGACAGGATTTCAACGCCAAATGCGGAGCCGTCACAGGGGCTCCAAAGACCAAGAG 720
QY 241 HisValValaIaIaThSerGlnProLySerThrcMetGlyHisGlnAapIyLysAapIleu 260
DB 721 CATGTGTAGCCACATCCACCAAGACCAATGCGGACATCAGGACAAAGAAAGATCATGT 780
QY 261 GlyGInCySerProLeuGlyCyLeuLeuIleArgIleLeuIleAapGlyIleSerLeuGly 280
DB 781 GGGGAGTGTCTCCCTTAGCTGCTCATCCGATATGATTGACGGATCTCTCCGG 840
QY 281 LeuThrcLeuArgLeuTPAapValTyrcLeuValGInGlyGInGlnAlaLeuMetProIle 300
DB 841 CTCACCTCGCGCTGTGGGACGTGATCTGGTAGAAGCGAAACAGCGCTTGTATGCCGATA 900
QY 301 ThrArgIleAlaPheLyValGInGlnLyAapArgLeuThrcIySerArgCyGlyPro 320
DB 901 ACMAAATCGCTTTAAGGTTCAGCAGAGGCGCTCAGAAAGACGTCCAGGTGTGCGCG 960
QY 321 TPAlaArgPheCyAspArgPheValAapThrTPAlaArgAapGluAapThrcValIleu 340
DB 961 TGGGACGTTTTTGCACCGGCTCTGTATACCTGGCCAGGACAGAGACATGTGCTC 1020
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrcArgIyGInGlyAapLeuProProPro 360
DB 1021 AAGCATCTTAGGGCTCTATGAGAAATCAACAAAGAGAGGAGGAGCTTGCAACCCCA 1080
QY 361 AlaLyPProGluGInGlySerSerAlaSerArgProValProAlaSerArgIyGlyLys 380
DB 1081 GCCAAACCCGAGCAAGGTCGTGCGCATCCAGGCTGTGCGGCTTCACTGCGGAGAG 1140
QY 381 ThrIleuCyLyArgIyAapArgGlnAlaProProGlyPProProAlaArgPheProArgPro 400
DB 1141 ACCCTCTGAGGGGAGACAGGAGGCCCCCTCCAGGCCACAGCCCGGTTCCCGGGCCC 1200
QY 401 IleTPSerAlaSerProProArgAlaProArgSerSerThrcProCyProGlyIyAla 420
DB 1201 ATTGTGACAGTTCCTCCGACAGGACCTGCTTCCACACCTGTCTGATGGAGGCT 1260
QY 421 ValaArgIuAapThrcTyrcProValGlyThrcGInGlyValProSerProAlaLeuAlaGln 440
DB 1261 GTCCGAGAAACACATCTACCTGTGGGCACTCAGGCTGTGCCAGCCGCTGCTCAG 1320
QY 441 GlyGlyProGInGlySerTPArgPheLeuGInTPAAsnSerMetProArgLeuProThrc 460
DB 1321 GAGGAGCTCCAGGGTCTCTGAGATCTCTGACAGTGAATCTCCATGCCCTCCCAAG 1380
QY 461 AapLeuAapValaIyGluGlyProTPPheArgHisTyrcAapPheArgIleSerCyArTPVal 480
DB 1381 GACCTGAGAGTAAAGAGGCTTGTGTCGACATTATGATTCAGACAGAGCTGTGGGTC 1440
QY 481 ArgAlaIleSerGInGluAapGlnLeuAlaProCyArTPGlnAlaGlnHisProAlaGln 500
DB 1441 CTTGCATATATCCAGAGAGACACAGCTGCGCCCTGCTGCGAGGCTGAACACCTGCGAG 1500

```

```

Oy 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 1501 CGGCTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACCAAGGAGACCCCTTCAGA 1560
Oy 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysGlyLeuHisLeu 540
Db 1561 GCTAGGAGACGACGAGCGTGTGCTCCACCTCAGGAGCTTGTGCTTGTGCGGCTCCTCAG 1620
Oy 541 GluSerSerGlnPheProGlyPhe 549
Db 1621 GAAGTTTCAGTTCCTCCAGGCTTC 1647

RESULT 3
AX775947 1964 bp mRNA linear PAT 14-JUL-2003
LOCUS AX775947
DEFINITION Sequence 217 from Patent WO03048202.
ACCESSION AX775947
VERSION AX775947.1 GI:32693665
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Matsuda, A. and Muramatsu, S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 217 12-JUN-2003;
AUTHOR Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
Source
1..1964
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1650
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAE11673.1"
/db_xref="GI:32693665"
translation="MDVVEVAGSWMAQEREDIMKYEKGRAGLPEDKGPFRSYNN
NVDHIGIVHETELPPLTAREAKQIRREISRSKRVMDLGMDEKTSRLIDRAVKG
PMNIRGPMWSTVLNIEMKLNKPGRYOIMKGRKSSSEHIQRIDRVSCTJAKHIFR
DRYGTQKRELIHLLAYEENPPEYCYRDSHIALFLYLPEBDAFVALVOLASER
HSLQGFSPNGVTQGLQDOEHVATISQRTWGHQDKDLGCGSPGLCLIRLIDG
DEPVLKHTASMKKLTQKQGLQPAKBOGSSASRPVPAISGGKTIKGRDRAAPR
PAPRPPRPIWASAPPRAPRSSTPCGAVAEEDTYPVGTGVPSPALACGCGSWRELU
QNMSPRLPTDLVVEGPFPRHYDFRQSCVWRAISODOLAPCMQAEHFAERKVSAPFA
PSTDSDGTFPRARDQPCAPISGPTCLGHLHSSQPPPGP"

ORIGIN

Alignment Scores:
Pred. No.: 2.23e-103 Length: 1964
Score: 2975.00 Matches: 548
Percent Similarity: 99.82% Conservatave: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.73% Indels: 0
Gaps: 0

US-10-071-838-2 (1-549) x AX775947 (1-1964)

Oy 1 MetAspValValGlnValAlaGlySerTrpTrpAlaGlnGlnuArgGlnuAspIleIleMet 20
Db 1 ATGCAAGCTGTAGAGCTCCGCGGAGTGTGTGGCAACAAGAGGAGCATCATTAAG 60
Oy 21 LysTrpGlnuArgGlyHisArgAlaGlyLeuProGlnuAspLysGlyProLysProPheArg 40
Db 61 AAATTCGAAAGGAGCAACGAGCTGGGCTGCAGAGGACAAAGGAGGCTTAAGCTTTTCA 120
Oy 41 SerTrpAsnAsnValAspHisLeuGlyTyrLeuHisGlnuThrGlnuLeuProLeu 60
Db 121 AGCTACAAACAACGCTGATCATTTTGGGATTGTATCATGAGCGAGCTGCTCTCTG 180
Oy 61 ThrAlaArgGlnuAlaLysGlnuIleArgArgGlnuIleSerArgLysSerIleTrpValAsp 80

181 ACTGGCGGAGGAGGAGCAAAATTCGCGGAGATCAGCGGAAGAGCAAGTGGATGAT 240
Oy 81 MetLeuGlnuAspTrpGlnuLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 300
Db 241 ATGCTGGAGACATGGGAGAAATACAAAGACACAGAAAGCTCATAGATCGAGCGCTACAG 300
Oy 101 GlyMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGlnuMet 420
Db 301 GGATGCCCATGAAACATCCGAGGACCTGATGCTCATCTCTGACATTAAGGAATG 360
Oy 121 LysLeuLysAsnProGlyLysArgTyrGlnuIleMetLysGlnuLysArgSerSerGlnu 140
Db 361 AAGTTAAAAAACCAGGAAGATACCATGATCATGAAGAGAGGCAAGAGCTCATCTGAG 420
Oy 141 HisIleGlnuArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 421 CACATCAGCGCATGACCGGAGAGTAAAGCGGACATTAAAGGAAGCATATATCTTCAG 480
Oy 161 AspArgTyrGlyThrLysGlnuArgLysLeuLeuHisIleLeuLeuAlaTyrGlnuLys 180
Db 481 GATCGATACGGAACCAAGCAGCGGAACTACTCCATCTCTCGCATATGAGAGAT 540
Oy 181 AsnProGlnuValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
Db 541 AACCGGAGGAGTGGCTACTGCGAGGACCTGAGGCACATCGCGCGCTTGTCTCTCTAT 600
Oy 201 LeuProGlnuLysAspAlaPheTrpAlaLeuValGlnuLeuAlaSerGlnuArgHisSer 220
Db 601 CTTCCTGAGAGGATGATCTGCGGACATGCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Oy 221 LeuGlnuLysPheHisSerProAsnGlyLysTrpValGlnuLysGlnuAspGlnuGlnu 240
Db 661 CTGAGAGGATTTCAAGCCCAATGGCGGACCTGAGCGGCTCCCAAGACCAACAGAG 720
Oy 241 HisValValAlaThrSerGlnuProLysThrMetGlyHisGlnuAspLysAspLeuCys 260
Db 721 CAGTGTGAGCCACGTCACACCAACCAAGACCATGGGCGATCGACCAAGAAATCTATGT 780
Oy 261 GlyGlnuCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 781 GGGCAGTGTTCCTCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Oy 281 LeuThrLeuArgLeuThrAspValTyrLeuValGlnuGlyGlnuAlaLeuMetProIle 300
Db 841 CTCACCTCGCGCTGTGGAGCTGTATCTGTGAAGGCGAAGCGCTTATGCCGATA 900
Oy 301 ThrArgIleAlaPheLysValGlnuGlnuLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 901 ACAGAAATCGCTTTAAGTTACAGAGAAAGCGCTTCACAGAAAGCTCCAGGTGGCGCG 960
Oy 321 TrpAlaArgPheCysAsnArgPheValAspTrpTrpAlaArgAspGlnuAspTrpValLeu 340
Db 961 TGGCAGCTTTTTCACACGCTGTGTGATACCTGGGCGAGGATGAGACATCTGTCTC 1020
Oy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnuArgLysLeuProPro 360
Db 1021 AAGCATCTTAAGGCTCTTATGAAGAAACTAACAAAGAACAGGGAGACTGCAACCCCA 1080
Oy 361 AlaLysProGlnuGlnuLysSerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1081 GCCAAACCGGACAGAGGCTGCTGGCATCCAGGCTGTGCGGCTTCAGTGGCGGGAAG 1140
Oy 381 ThrLeuCysLysGlyAspArgGlnuAlaProProGlnuLysProProAlaArgPheProArg 400
Db 1141 ACCCTGTGAAAGGAGGACAGAGCGCTTCAGGCCACCAAGCCGAGTCCGCGCGCC 1200
Oy 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysAspGlnuGlyAla 420
Db 1201 ATTGTGTCAGTTCCTCCCGGACAGGACCTGCTTTCACACCTGCTGCTGCTGCTGCT 1260
Oy 421 ValArgGlnuAspTrpTyrProValGlyThrGlnuGlyValProSerProAlaLeuAlaGln 440

```

Df		1261	GTCGGGGAAGACCTTACCCTGTGTGGGCACTCAGGGGTGTGCCAAGCCGGCCCTGGACTG	1326
OY		441	GLYGIYProGINGLYSerTPHArgPheLeuGINTPanSerMetProAlaLeuProThr	460
Df		1321	GGAGGACCTCAGGGGTTCCTCGAGAATCTCTGCACGTGGAACTCCATGCCGCCCTCCCAAG	1380
OY		461	AAPLeuAPValIGluGIYProTPHArgRHSIETPAAPPhArgGINSerCYeTPVal	480
Df		1361	GACCTGGACGTGAGGGGCCCTTGTTGCTGGCATTAATTCAGACAGAGCTGTGGGTTC	1440
OY		481	ARGAlalieserGINGLUAAPGINleuAlaproCYSTRIInalAGLNIAPAlaGlu	500
Df		1441	CCTGCATATCCCAAGAGACACAGCTGGCCCCCTCTGTCGACGGCTGAAACACCTGTGGAG	1500
OY		501	ArgValArgSerAlaphAlaAlaProSerThrapSerArpGINLYTHProPhArg	520
Df		1501	CGGGTAGAATCGGCTTTCGTGTGCACCCAGCACTGATTCCGACCAAGGGCACCCCCTTCAGA	1560
OY		521	AlAArgArpGIUGINProCYAlaAProPhArgGIYPRoCYleuCyGlylleuHIleu	540
Df		1561	GCTAGGGAGCAACAGCCGTGTGTCTCCCACTCAGGGCCCTTCCTCTGTGGCCCTCCACTTG	1620
OY		541	GlusEserGINphProProGIYPHe	549
Df		1621	GAAGCTTCTCAGTTCCTCCCTCCAGGCTTC	1647
RESULT 4				
LOCUS	BC071680	2090 bp	mRNA	linear PRI 30-JUN-2004
DEFINITION	Homo sapiens TBCL domain family, member 3, mRNA (cDNA clone			
ACCESSION	MGC:87851 IMAGE:5165385), complete cds.			
VERSION	BC071680			
KEYWORDS	BC071680.1 GI:47939522			
SOURCE	MGC.			
ORGANISM	Homo sapiens (human)			
REFERENCE	Bukaryotai Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 2090)			
	Straubeberg R.L., Collins,F.D., Wagner L.H., Dergs,J.G.,			
	Klausner R.D., Collins,P.S., Moore,T., Max,S.I., Wang,J.J., Helen,F.,			
	Alschul S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhac,N.K.,			
	Hopkins,R.F., Jordan,H., Moore,T., Farmer,A.A., Rubin,G.M., Hong,L.,			
	Ditachenko,L., Marusina,K., Palmer,A.A., Rablin,G.M., Hong,L.,			
	Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,			
	Scheer,T.E., Brownstein,M.J., Udell,T.B., Toshitsuki,S.,			
	Carinci P., Prange C., Raha S.S., Loquellano N.A., Peters,G.J.,			
	Abramson R.D., Mullenhy,S.J., Bosak,S.A., McEwan,P.J.,			
	McKernan K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,			
	Woley,K.C., Hale,S., Garcia,A.M., Gay,L.U., Hulye,S.W.,			
	Villallon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,			
	Fahney,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,			
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,			
	Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,			
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,U., Myers,R.M.,			
	Butterfield,V.S., Krzyzinski,I., Mi, Skalska,U., Smilins,D.E.,			
	Schenerich,A., Schein,J.R., Jones,S.J. and Merritt,M.A.			
TITLE	Generation and initial analysis of more than 15,000 full-length			
JOURNAL	human and mouse cDNA sequences			
PUBMED	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
REFERENCES	12477932			
AUTHORS	2 (bases 1 to 2090)			
TITLE	Straubeberg,R.			
JOURNAL	Direct Submission			
REMARK	Submitted (01-JUN-2004) National Institutes of Health, Mammalian			
COMMENT	Gene Collection (MGC), Cancer Genomics Office, National Cancer			
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
	USA			
	NIR-MGC Project URL: http://mgc.ncl.nih.gov			
	Contact: MGC help desk			
	Email: cgapsb-remail.nih.gov			
	Tissue Procurement: Life Technologies, Inc.			
	cDNA Library Preparation: Life Technologies, Inc.			

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@pdx1.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNT at: <http://image.llnl.gov>
Series: IRAK Plate: 166 Row: g Column: 13
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14145984.

FEATURES

Source

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.:	3,35e-103	Length:	2099
Score:	2371.00	Matches:	546
Percent Similarity:	99.83%	Conservative:	2
Best Local Similarity:	99.45%	Mismatches:	1
Query Match:	99.60%	Indels:	0
DB:	9	Gaps:	0

US-10-071-838-2 (1-549) x BC071680 (1-2090)

QY	MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluAlaGlyIleLeuMet	20
Db	ATGAGAGCTGTAGAGGTCCGGGCAAGTGGTGCGCACAGAGCCAGAGACATCATTTATG	160
QY	LYeTYrGluTyAGlyVHIsaRglaGlyLeuProGluAspLysGlyProLysProPheArg	40
Db	AAATTAGAAAAGAGACCCGAGCTGGGTGCACAGAGCACAGGGGCTTAAGCCTTTTCCA	220
QY	SeTYrAsnaAsnaValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu	60
Db	AGCTACAAACAACAGCTGCATCATTTGGGAGTTTGACATGACGAGAGCTGCCCTCTCG	280
QY	ThrAlaArgGluAlaGlyGlnLeuArgArgGluIleLeuSerArgLysSerLysTrpValAsp	80
Db	ACTGCCCGGAGGCGCAACAAATTTGGGGGAGAGTACGCCGAAAGAGCAGTGGGTGAT	340

QY 81 MetLeuGIYAAPTTrGluLysTyrIleSerSerValIleuIleAspArgAlaTyrLys 100
 Db 341 ATGCTGGAGACGTGGAGAAATACAAAGACGAGAAAGCTCATAGATCGAGGTACAG 400
 QY 101 GlyMetProMetAsnIleArgIleProMetTrpSerValIleuIleuennileuIleuMet 120
 Db 401 GGAATGCCCATGACATCCGGGGCCGATGTGTCACTCTCTGAAACATTTGAGGAAAG 460
 QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetIleGlyLysGlyLysArgSerSerGlu 140
 Db 461 AAGTTGAAAAACCCCGGAAATACCAATCATAGAGAGAGGCAAGGATCATCTGAG 520
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 521 CACATCCAGCGCATCGACCGGACATAGCGGACATTTAAGAAAGCATATGTTCTTCAGG 580
 QY 161 AspArgTyrGlyThrIleLysGlnArgGlyLysLeuLysHisIleLeuLeuAlaTyrGluLysTyr 180
 Db 581 GATCGATACGAAACCAAGACGCGGAACTACCTCCATCTCTCTGCAATAGAGAGTAT 640
 QY 181 AsnProGlyValGlyTyrCysArgAspLeuSerHisIleAlaIlePheLeuLeuTyr 200
 Db 641 AACCCGAGAGTGGGTACTGACGAGGACCTGACCCACATGCGCCTTGTCTCTCTAT 700
 QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 701 CTTCCTGAGAGAGATGATCTTGGGACCTGTGTGACGTGCTGGCCAGTGAAGGACATCC 760
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
 Db 761 CTGCGAGGATTTCAAGCCCAATGGCGGACCTCCAGGGGCTCCAAAGCCAAAGAGAG 820
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
 Db 821 CATGTGTAGACGATCGTCAACCAAGCCATGGGAGCATCAGAGCAAGAAAGTCTATGT 880
 QY 261 GlyIleCysSerProLeuGlyCysLeuIleArgIleLeuIleArgGlyIleSerLeuGly 280
 Db 881 GGGCAGTGTTCCTGTTAGGCTGCTCATCCGATATTTGATACGAGGATCTCTCTCGGG 940
 QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnIleLeuMetProIle 300
 Db 941 CTCACCTCGCCCTGTGTGAGCGTGTATCTGTGAAAGCGAAGCAGCGTTGATGCCGAT 1000
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgCysGlyPro 320
 Db 1001 ACAAGATCCGCTTTAAGTTGACAGAGCCGCTCAAGAGCGTCCAGGTGTGCGCG 1060
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 1061 TGGGCAAGTTTGGCAACCGGTTGTTGATACCTGGGCGAGGATGAGGACATGTGCTC 1120
 QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
 Db 1121 AAGCATCTTAGGGCTCTATGAAGAAACCTAACAAAGACAGGGGACCTCCACCCCA 1180
 QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
 Db 1181 GCCAAACCCGAGACAGAGGCTGTGCGCATCGAGCTGTGCGGCTTACGTTGCGGGAAG 1240
 QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1241 ACCCTCTGCAAGGGGACAGGAGGCCCTCCAGAGCCCAAGGCCGAGTTCCCGCGGCC 1300
 QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
 Db 1301 ATTGTGATGCTTCCCGCGACGAGGACATCTGTTCTTCCACACCTGTCTGTGTTGGGCT 1360
 QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValaProSerProAlaLeuAlaGln 440
 Db 1361 GTCCGGGAAAGACCTACCTGTGTGGCATCTAGGGTGTGCGAGCCCGGCTGTGCTGAG 1420

QY 441 GlyIleProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
 Db 1421 GGAAGACCTCAGGGTCTCTGGAATCTCTGCAATGAGTCCATGAGCCCGCTCCCAAG 1480
 QY 461 AsnLeuArgValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
 Db 1481 GACCTGAGCGTGAAGGAGCCCTTGTGTCCCGCATTTATGATTTCAAGACAGGCTGCGGTC 1540
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
 Db 1541 CGTGCATATCCAGAGAGACACAGCTGCGCCCTCTCTGCAAGGCTGAACACCTGCGAG 1600
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 Db 1601 CGGTTAGATCGGCTTTGCTGTGACCCAGCACTGATTCGACAGGGGACCCCTTCAGA 1660
 QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
 Db 1661 GCTAGGAGCAACACAGATGTGCTCCACCTCAGGGGCTTGTGCTGCGGCTTCACCTTG 1720
 QY 541 GluSerSerGlnPheProProGlyPhe 549
 Db 1721 GAAAGTCTCAGTCTCTCCAGGCTTC 1747
 RESULT 5
 AX086847
 LOCUS AX086847 2072 bp DNA linear PAT 09-MAR-2001
 DEFINITION Sequence 799 from Patent WO0112659.
 ACCESSION AX086847
 VERSION AX086847.1 GI:13276047
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1. Mleemann, S. Human data sequences
 AUTHORS TITLE
 JOURNAL Human Genome Project (DB)
 FEATRES Location/Qualifiers
 SOURCE 1. 2072
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,63e-103 Length: 2072
 Score: 2963.00 Matches: 545
 Percent Similarity: 99.64% Conservative: 2
 Best Local Similarity: 99.27% Mismatch: 2
 Query Match: 99.33% Indels: 0
 DB: 6 Gaps: 0
 US-10-071-838-2 (1-549) x AX086847 (1-2072)
 QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
 Db 99 ATGACGATGTAGAGGATCGCGGCACTTGTGGGACACAAAGGAGGACATCATTTAG 158
 QY 21 LysTyrGlyLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 Db 159 AAATACGAAAAGGACACCGAGCTGGGCTGCAGAGGACAAAGGCGCTTAAGCTTTTGA 218
 QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGlyIleThrGluLeuProLeu 60
 Db 219 AGTACAAACAACAACGTCAATCTTTGGGATTTGATCATGAAACGAGACTGCTCTCTG 278
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerIleTrpValAsp 80
 Db 279 ACTGCGGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGACAGATGGGTGAT 338

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 2166)
 AUTHORS Struhsberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Logucliano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M., Butlerfield, S., Krzywinski, M.I., Skalska, U., Smal, D.B., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL 12477932
 PUBMED 2 (bases 1 to 2166)
 REFERENCE Director MGC Project.
 TITLE Direct Submission
 AUTHORS Submitted (06-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 COMMENT NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 CONTACT: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-ehg.stanford.edu>
 Contact: (Dickson, Mark) mcg@axll.stanford.edu
 Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 166 Row: h Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.
 Location/Qualifiers
 1. 2166
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:87892 IMAGE:5170064"
 /tissue_type="Brain, adult medulla"
 /clone_1ib="NIH MGC_119"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. 2166
 /gene="LOC414059"
 /db_xref="LOC414059"
 147. 1796
 /gene="LOC414059"
 /codon_start=1
 /product="centromeric TBC1 domain family member 3-like"
 /protein_id="AAH75809.1"
 /db_xref="GI:43902677"
 /db_xref="LOC414059"
 /translation="MDVVAVAGSMMVAQEREDIIIMYKKGRAGLPEDKPKPRSYNN
 NVNDLGIYHETELPLTAREAKQIRREISRSKSKWMDLGDWEKYSRRLIDRAVKGM

ORIGIN
 Alignment Scores:
 Pred. No.: 3,276-102 Length: 2166
 Score: 2945.00 Matches: 541
 Percent Similarity: 99.27% Conservative: 4
 Best Local Similarity: 98.54% Mismatches: 4
 Query Match: 98.73% Indels: 0
 DB: 9 Gaps: 0
 US-10-071-838-2 (1-549) x BC075809 (1-2166)
 1 Metaaapvaialgluvalaiaaglyserpttpalaglngluargluaspillemet 20
 147 ATGACGTGTGTAGAGGTCCGGGTGTTGTGTGGCAACAAGCAGAGACATCTATG 206
 21 LysTYRGluLYeGLYHiaRgaiaGLYLeuProGLuaapLYeGLYProLYsProPhearg 40
 207 AAATACGAAAGGACACCGAGCTGGCTCCAGAGCAAGGGCCTTAAGCTTTTCA 266
 41 SerTYAsnaaanaValaaphiaLeunglyilevalhagluThgluleuProPoleu 60
 267 ACTCAACAACAACATGTCATCTTTGGGATTTGATGAGACGAGCTGCTCTCG 326
 61 ThrAlaahrgluAlaalyglnleahrgarvgulileserArgLYserLYerTYValaap 80
 327 ACTGCGCGGAGGACCAAGCAATTCGGCGGAGATCAGCGAAGCAAGTGGTGGAT 386
 81 MetLeuGLYAspTYRGLYLYserLYsSerSerArgLYLeuLeuAspArgLaTYLYs 100
 387 ATGCTGGAGATCGGAGAAATACAAAGACCAAACTCATATGATGACCTACAG 446
 101 GlyMetProMetAsnIleArgLYProMetTYPSerValLeuLeuAsnIleGLuIuMet 120
 447 GGAAATGCCATGAATCATCGGGGCCGATGTGTGATCTCTCGAATGAGCAAAATG 506
 121 LysLeuLYAsnProGLYArgTYRglnIleMetLYsGLuLYsGLYLYArgSerSerGLu 140
 507 AAGTTGAAAAACCCCGAAGATACCATGAGAGAGAGAGAGGATCATCTGAG 566
 141 HisIleGLNArgIleAspArgAspValSerGLYThrLeuArgLYHisIlePhePhearg 160
 567 CACATCCAGCGCATGACCGGACATTAAGCGGACATTAAGGACATATGTTCTTCAAG 626
 161 AspArgTYRGLYThLYsGLNArgLYLeuLeuHisIleLeuLeuAlaTYRGLuIuTYR 180
 627 GATTCATACGGAACCAAGAGCGGAGATCATCTCTCTCGCATATAGAGATAT 686
 181 AsnProGLuValGLYTYRcyArgAspLeuSerHisIleAlaAlaLeuPheLeuTYR 200
 687 AACCCGAGGAGGTGTACTGACGAGACCTGAGCCATCGCCGCTTGTCTCTCTCAT 746
 201 LeuProGLuGLuAspAlaPheTYRAlaLeuValGlnLeuLeuAlaSerGLuArgHisSer 220
 747 CTTCCTGAGAGAGATGATCTTGGGACATGATGAGCTGTGGCCATGAGAGCATCTC 806
 221 LeuGLNArgLYPheHisSerProAsnGLYGLYThValGlnGLYLeuGlnAspGLNGLu 240
 807 CTGCAAGGATTTCAAGCCCAATGCGGAGACCGTCCAGGGGCTCCAAAGCAAGAG 866
 241 HisValValAlaThSerGLNProLYserThMetGLYHisIleAspLYLYsAspLeuLYs 260
 867 CATGTGTAGCCATCATCAATCAAGACATGAGGAGCATCGAAGAAAGATCTATGT 926
 261 GlyGlnCYsSerProLeuGLYCYsLeuIleArgIleLeuIleAspGLYIleSerLeuGLY 280


```

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 962 CTTCCTGAGAGAGATGCACTTCTGGGCACTGGTGCACCTGCTCCAGTACAGGCACTTC 1021
QY 221 LeuGlnGlyPheHisSerProAnglyGlyThrValGlnGlyLeuGlnAspGlnGlnGly 240
DB 1022 CTGCAAGGATTTTCAACAGCCCAAAATGGCGGAGCGTCCAGGGGCTCCAAACCAACAGGAG 1081
QY 241 HisValValAlaThrSerGlnProGlyThrMetGlyHisGlnAspGlyLeuAspLeuGly 260
DB 1082 CATGGTGAAGCAGCATCCACMACCCAGACCATGGGGCATCAGAGCAAGAAATCATATGT 1141
QY 261 GlyGlnGlySerProLeuGlyCysLeuLeuLeuArgGlnLeuAlaAspGlyIleSerLeuGly 280
DB 1142 GGGCAGTGTTCCTCCGTTAGGCTGCTCATCCGATTTGATTTGAGGGATCTCTCTCGGG 1201
QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
DB 1202 CTCACCTCGCGCTGTGGGACGTGATCTGTAGAAAGGCAACAGGCGTTGATGCCGATA 1261
QY 301 ThrArgGlnLeuAlaPheLeuValGlnGlnLeuValGlnLeuThrLeuThrSerArgGlyPro 320
DB 1262 ACMAAATCGCCTTTAAGGTTTCAAGCAGAGCGCTTCACAGACAGCTCCAGGTGTGGCCG 1321
QY 321 TrpAlaArgPheCysAsnAspPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 1322 TGGGACCGTTTGTGCAACCGGTTCTGTATACCTGGGACGAGGATGAGACCTGTGCTC 1381
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
DB 1382 AAGCATCTTAAAGGCTCTATGAAAGAACCAAGAAACAGAGGAGGAGCTGCACCCCA 1441
QY 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgLysGlyLys 380
DB 1442 GCCAAACCCGAGCAAGGCTGTGGGATCCAGGCTGCTCCGCTTCACGTGCGGAG 1501
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1502 ACCCTCTGAGAGGAGACAGGAGGCGCTCCAGG-----CCA 1540
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1541 TT-TGTGACGCTCCCGGCAAGGAGCCTGCTTCTTCCACACCTGCTCTGTGGGCT 1599
QY 421 ValArgGluAspThrTyrProValGlnGlnGlnValProSerProAlaLeuAlaGln 440
DB 1600 GTCCGGAGAGACACTTACCTGTGGGCACTCAGAGGTGTCCAGCCCGGCTCGAG 1659
QY 441 GlyGlyProGlnGlySerThrPhePheLeuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1660 GAGAGACCTCAGAGGTTCTGAGATTCCTGACGTGAGACTCCATGCCCCGCTCCCAACG 1719
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1720 GACCTGGAGAGTGAAGGCGCTGTTCCGCAATTGATTTCAAGACAGAGCTGCTGGGCT 1779
QY 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
DB 1780 CGTGCATATCCCAAGAGAGCAGCTGCGCCCTGCTGAGGCTGAAACACCTCGGAG 1839
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1840 CGGGTGAAGTGGCTTTCGCTCACCAGACCTGATTCGACACAGGAGCCCCCTTCACA 1899
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLysHisLeu 540
DB 1900 GCTAAGAGAGAGACGCGTGTCTCCACTCAGAGGCTTGTGCTCTGCGGCTCCACTTG 1959
QY 541 GluSerSerGlnPheProGlyPhe 549
DB 1960 GAAAGTTTCAGTTCCTCCAGGCTTC 1986

```

```

AX775941
LOCUS AX775941 2304 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 211 from Patent WO03048202.
ACCESSION AX775941
VERSION AX775941.1 GI:32693659
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Matsumura, A. and Muramatsu, S.
NF-KB activating gene
Patent: WO 03048202-A 211 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
source
1..2304
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
362..1732
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11670.1"
/db_xref="GI:32693660"
/translation="MDVYVAGSMWAQREDDIMKYEKGRAGLPEDKGPKEPSYNN
NDHIGIYHEBELPRLTAREAKQIAREISRSKWDMDGMEKYSRKLIDRAYKM
PMNIRGPMVSVALNTBENKLNKPGYQIMKEKGRSSSHQRIIDVDVSGTLRKHIFPR
DRYGTQRELLHLILAYEENPEVGYCDLSHIALFLYLPEDBAFALVOLLASER
HSLQGFHSFNGSTVQGLDQOEHVAVATQPTMGQDKDKDLCGQSPGLRLIIDG
ISLGLTLFLMPVYLVEGEOALMPITRIAFKQOQKLTYSKCGPWARCNRPVDMAR
DEDVTLKHLRASMKLTLRKQGLDLPAPAEQSSASRPVPSRGGKTLCKGRDAPPG
PBGQIPRGLHVLPHPVLVGLSGKPTLMALRYCPARWLRBDRLPEDSCSGTCEPA
SORTWT"
ORIGIN
Alignment Scores:
Pred. No.: 1,926-100 Length: 2304
Score: 2898.50 Matches: 539
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 3
Query Match: 97.17% Indels: 8
DB: 6 Gaps: 1
US-10-071-838-2 (1-549) x AX775941 (1-2304)
QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet 20
DB 362 ATGGAAGTGTGAGAGGTGCGGGGAGTGTGTGGGCAACAAAGCAGAGGACATCATTAATG 421
QY 21 LysTyrGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 422 AAATCGAAAGGACACCGAGCTGGCTGCCAGAGGACMAAGGCGCTTAAGCCTTTTCCA 481
QY 41 SerTyrAsnAspAspValAspHisLeuGlyIleValHisGlnThrGlnLeuProProLeu 60
DB 482 AGCTACACACACAGCTGCATCATTTGGGAGATTGTATCATGAGACGAGCTGCTCTCG 541
QY 481 ThrAlaArgGluAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80
DB 542 ACTGGCGGAGGCGCAAGCAATTCGGCGGAGATCAAGCGCAAGAGCAAGGTGGTGAAT 601
QY 81 MetLeuGlyAspTrpGlnLysTyrLysSerSerArgLysLysLeuIleAspArgAlaTyrLys 100
DB 602 ATGCTGGAGAGATCGGAGAAATACAAAGCAGCAAGCAAGAGCTCATGATGAGCGTACAG 661
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnIleMet 120
DB 662 GGAATGCCCATGAACATCCGGGAGCCGATGTGTCTGATGCTCTCTTAACCATGAGAAATG 721
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGln 140
DB 722 AAGTTGAAAAACCCCGAAGATACCATGATCATGAGAGAGAGGAGGAGGTGATCTGAG 781

```

[illegible]

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLAST at: <http://image.jnl.gov>
 Series: IRAX Plate: 68 Row: 1 Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 14149984.

FEATURES

source

1. 2575
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:44903 IMAGE:5167394"
 /tissue_type="Brain, adult medulla"
 /clone_1db="NIH_MGC_119"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1. 2575
 /gene="MGC44903"
 /db_xref="locusID:414060"
 78. 1208
 /gene="MGC44903"
 /codon_start=1
 /product="similar to TBC1 domain family, member 3"
 /protein_id="AAH33670.1"
 /db_xref="gi:45708751"
 /db_xref="locusID:414060"
 /translation="MDVAVAGSSWMAQEREDILIMKYEKIRAGLPEDKPKPPRSYNN
 NVDLGIIVHETELPLTAEAKQIRREISRSKSWDMGLDWKRYSSRLIDRAKGM
 PNRKGPWMSVLNIEEMKKNPGRYQIMKEGKSSSEIQRIDIDISGLTKHMFPR
 DRYGKREBLHILAEYENVEYCRDLSHIALFLYLEPDEAFMLVQLASER
 HSLGFHSNGSTVGLODOOHVAVTSKTMGKDKDLGCGSPICGLIRLIDG
 VRRHETLAQGSPLPCSAALRPGGLAHSPPTGSGSGSPKDOAPASQESSQGNWG
 GSGPMSMDMNRTRLEKQSLKPWASPSLQEPHPSNECP"

gene

CDS

1. 54e-98 Length: 2575
 Score: 2849.00 Matches: 543
 Percent Similarity: 77.41% Conservative: 2
 Best Local Similarity: 77.13% Mismatches: 4
 Query Match: 95.51% Indels: 155
 DB: 9 Gaps: 1

ORIGIN

Alignment Scores:

Pred. No.: 1.54e-98 Length: 2575
 Score: 2849.00 Matches: 543
 Percent Similarity: 77.41% Conservative: 2
 Best Local Similarity: 77.13% Mismatches: 4
 Query Match: 95.51% Indels: 155
 DB: 9 Gaps: 1

US-10-071-838-2 (1-549) x BC033670 (1-2575)

QY 1 MetAepValValGluValAlaGlySerTPTPAlaGlnGluArgGluAspIleIleMet 20
 DB 78 ATGACGTGTAGAGGTGCGGGGTAGTTGTTGGGCAAGAGGAGAGACATCATTTATG 137
 QY 21 LysTyrGluValGlyYHIAArgAlaGlyLeuProGluAspGlyProLysProPheArg 40
 DB 138 AATATCGAAGAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCA 197
 QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
 DB 198 AGCTTCAACAAACAGCTCATCTTTGGGATTGTTCATGAGACGAGCTGCTCTCTG 257
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTyrValAsp 80
 DB 258 ACTGCGCGGAGCGCAAGCAAAATTCGCGGAGATCAAGCAAGAGCAAGTGGTGAT 317
 QY 81 MetLeuGlyAspTyrGluValGlyYHIAArgAlaGlyLeuProGluAspGlyProLys 100
 DB 318 ATGCTGGGAGAGTGGAGAAATACAAAGACAGCAAGAACTCATATGAGCGTACAG 377
 QY 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
 DB 378 GGAATGCCCATGAACATCCGGGGCCGAGTGTGCTGCTCTGTAACATTAAGAAAG 437
 QY 121 LysLeuValAsnProGlyValArgTyrGlnIleMetLysGluValGlyLysArgSerSerGlu 140
 DB 438 AAGTGAAGAAACCCCGAGATACCAAGATCATGAAGAGAAAGGCAAGAGTCACTGAG 497

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 DB 498 CACATTCACGCGCATGACCGGACATTAAGCGGACATTAAGAGATATGTTCTTCAG 557
 QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
 DB 558 GATCATACGGAACCAAGACGCGGAACTACCTCCATCTCTCTGCAATATAGAGATAT 617
 QY 181 AsnProGluValGlyTyrCysArgArgSerSerHisIleAlaIleAspPheLeuTyr 200
 DB 618 AACCCGAGAGTGGCTACTGACAGGACCTGAGCCACATGCGCGCTGTCTCTCTAT 677
 QY 201 LeuProGluGluAspAlaPheThrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 DB 678 CTTCCTGAGAGAGATGATCTGAGGACCTGGGACAGCTGCTGCGCAAGTGAAGGACCTCC 737
 QY 221 LeuGlnGlyPheHisSerProLeuGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 DB 738 CTGACAGGATTTACAGCCCAATGCGGACCGTCCAGGGGCTCCAGACCAAGAGAG 797
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
 DB 798 CATGTGTAGCCACGTCACATCAAGACCATGAGGACATGAGACAAAGATCATATGT 857
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGly----- 276
 DB 858 GGGCAGATGTTCCCGTTAGCTGCTCATCCGGAATTTGATTGACGGGGTAAGAGGCAT 917
 QY 276 ----- 276
 DB 918 AGGAGAACCCCTGCTCAGGAGACCTTCCTTGCCCTGACAGTCCCTCTCCAGCCCGG 977
 QY 276 ----- 276
 DB 978 GGTCTGCTACTCCAGCCCAAGAGGCTCAGCGGGTCCCAAGACACACAGCA 1037
 QY 276 ----- 276
 DB 1038 AAACCTCTGCCCAAGAGGGGTCTATCCAGGCAATGCTGGGGCTCAGGCCAGCTCA 1097
 QY 276 ----- 276
 DB 1098 TGGGACAGCTGGGCGCAGGACCCGACTGAGAGGCTCAGGAGAAAGCTCAAGCCTTGGGCA 1157
 QY 276 ----- 276
 DB 1158 AGCCCTCTCTCAGAGACCATCCCACTCAATGATGCTCCCATGAGAGACTTCA 1217
 QY 276 ----- 276
 DB 1218 AGACTTGTCTGACCCAGCGTCTGAGAGGCTCAGCCGACCTCATGGGAAAGTTCATCG 1277
 QY 276 ----- 276
 DB 1278 ACTGTGAGAGACTGAAGCCCGCAGCTGCGAGCTGAGCCACAGCCCGAGCTGAGAGA 1337
 QY 277 ----- 277
 DB 1338 CCAAGTCTTTTCACTGCTGTCTCCACAGATCTCTCTCGGCTTCACCTCGCTG 1397
 QY 286 PAspValTyrLeuValGlnGlyGluGlnAlaLeuMetProIleThrArgIleAlaPheLys 306
 DB 1398 GAGCTGTCTGTGTAGAGGCAACAGCGTGTATGCCGATACAGAAATGCGCTTTTA 1457
 QY 306 ValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTPAlaArgPheCysAs 326
 DB 1458 GGTTCAGAGAGAGCCCTTCACAGAACGTCAGTGTGGCCGCGTGGGCACTTTTTCGAA 1517
 QY 326 NArgPheValAspThrTPAlaArgAspGluAspThrValLeuLysHisIleValAlaSer 346
 DB 1518 CCGGTTCTGTATACCTGGGCGAGGATGAGGACCTGTGCTCAAGACATCTTAAGGCTTC 1577
 QY 346 rMetLysLysLeuThrArgLysGlnGlyAspLeuProProAlaLysProGluGlnGln 366

```

Db      1578 TATGAGAACTAACAAGAAAGAGGGGAGCTGCAACCCCAACCAACCCCAAGG 1637
Qy      366 YSERSERAAASERATPROVALPROALASERAGLYGLYVSTHLEUCYSLYSEGLYAS 386
Db      1638 GTGTGTGGGATCCAGGCTGTGTGGCTTCAACGAGGGGAGAACCTCTGCAAGGGGGA 1697
Qy      386 PARGSLNALAPROGILYPROGILAAATGPHETPRAATGPRGILLETIPSERALASERPR 406
Db      1698 CAGGCAAGCCCTCCAGGCCCCAGCCAGCCGCTTCCGCGGCCCAATTGGTCAAGCTTCCCT 1757
Qy      406 OPRATPAPALAPROAGSERSETHPRCYSPROGILYGLYVALAVALAGLUAAPRTHY 426
Db      1758 GCCACGGGACCTGTGTCTTCCACACCTGTCTGTGTGGGCTGTGTGGGAGACACCTTA 1817
Qy      426 RPROVALGLYTHGNGLYVALPROSERPROALALEUALAGLNGLYGLYPROGNGLYSE 446
Db      1818 CCGTGTGGGCACTCAGGGGTGTGCCAGCCGCGCTGTGCTCAGGAGAGACCTCAGGGTTC 1877
Qy      446 RTTPATPAPHELEUNGINTTPASMSERMETPROATGLEUPROTHRAPLEUAPVALGLUGL 466
Db      1878 CTGAGATTCCTGCACTGAGTGAACCTCCATCCCGCTCCCAACGSACTTGAAGAGG 1937
Qy      466 YPROTPPHEATGHSITYRASPHEATGINSERCTTPVALARGALIASERGLNGI 486
Db      1938 CCGTGTGTGCTCCCATATATGATTTCAACAGAGCTGTGTGTGCTGTGTGCAATCCAGGA 1997
Qy      486 UASPOLLEUALAPROCYSTTPGINALAGLUHSIPROALAGLUARGVALARGSERALAPH 506
Db      1998 GGACCAAGCTGGCCCCCTGTGTGTGAGAGCTGAACCCCTGCGAGCGGGGTGAGTCCGCTTT 2057
Qy      506 EALALAPROSERTHRAPSERASAPGINGLYTHRPHAPHEARGALAAARGAPGLUGINPR 526
Db      2058 CACTGCACCAAGCACTGATATCCAGACAGGCAACCCCTTCAAGCTAGAGAGCAAGACGCA 2117
Qy      526 OCYALAPROTHSERGLYPROCYALEUCYSEGLYLUHSILEUNGILUSERSERGINPHEPR 546
Db      2118 GTGTGTCCCACTCAGGCGCTTGGCTTGTGCGGCTTCCACTTGGAAGATTCTCAGTTCCC 2177
Qy      546 OPROGLYPHE 549
Db      2178 TCCAGGCTTC 2187

RESULT 11
LOCUS   AX775951 2647 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 221 from Patent WO03048202.
ACCESSION AX775951
VERSION  AX775951.1 GI:32693669
KEYWORDS
SOURCE  Homo sapiens (human)
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Matsuda, A. and Muramatsu, S.
AUTHORS   NF-KB activating gene
TITLE     Patent: WO 03048202-A 221 12-JUN-2003;
JOURNAL   Asahi Kasei Kabushiki Kaisha (JP)
FEATURES  Location/Qualifiers
            source          1..2647
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /note="unamed protein product"
                        /codon_start=1
                        /protein_id="CA811675.1"
                        /db_xref="GI:32693670"
                        /translation="MDVVEVAGSWAQEREDIIIMYKEXHRAGLPEKGPPEPSYNN
                        PMNIRGPMWVILNTEEMKLNKRGYQIKKKGKSSSHIQRIDRVSGTLRKHIFPR
                        DRIGTKORELHITLAEYENPEVGCNDLSHIALFLILYPERDARWALVQLASER

```

```

HSLGAFHSBNGSTVQGLIODOEHVAVTSQPKTWGHOISGLTLRLMDVLYVREGSALM
PITRAAPFVQOKRLITKTSRCGPWAFRCNPFVDPWADDEDTVTKHILRAKMKUTRQGD
LPPAKPQSGSSASRPVPAASRGKTLCKSDROAPPEPPAPRPRPIMASPPRARSST
PCRGATVREDTYPVGTGVPSPALACGGGCGSMRPLQMSNRPRLTDLVDGPFWRHY
DPKSCVWRAIISQEDLAPCWMQEHNPABERVSRAPAAPSTDSQGTFFRARDPCAPLT
SGPCLGLHLESSQFPFPGF"

```

ALIGNMENT SCORES:

```

Score: 5.3e-98 Length: 2647
Pred. No.: 2835.00 Matches: 526
Percent Similarity: 95.81% Conservative: 0
Best Local Similarity: 95.81% Mismatches: 1
Query Match: 95.04% Indels: 22
Gaps: 1

```

US-10-071-838-2 (1-549) x AX775951 (1-2647)

```

Qy      1 MetASPVALVALGLUVALIAGLYSERTRTPALAGLNGLUARGLUASPILEIEMET 20
Db      755 ATGACGCTGTAGAGCTGCGGCGAGTGTGGGCAACAGAGCGAGGACATCTATTG 814
Qy      21 LYSYRGLULYSGLYHISATGALAGLYLEUPROGLUASPLYSGLYPROLYSPROPHET 40
Db      815 AAATACGAAAGGAGACACGAGCTGGGCTGCGAGAGACAAAGGGCTTAAGCTTTTTCGA 874
Qy      41 SERTYRANANANVALAEPHISLEUNGILYILEVALHISGLVTHGLULEUPROPROLEU 60
Db      875 AGCTAACAAACAACGTCATCTTTGGGATTTGACATGAGACGAGCTCTCTCTG 934
Qy      61 THR1AARGSLUALYSGNLINLEARGILULEARGILULESERATGLYSESTYRTPVALASP 80
Db      935 ACTGCGGGAGGCGCAAGCAAAATTCGCGGAGATCAGCCGAAAGACCAAGTGGTGGAT 994
Qy      81 METLEUGLYASPTTGLULYSTYRYSERSESTARGLYLEULEAPATGALATYRGLYS 100
Db      995 ATCTGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCAATGATCGAGCTACAG 1054
Qy      101 GLYMETPROMETASNL1ARGGLYPROMETTSPSERVALLEULEUAN1LEGLUGIMET 120
Db      1055 GGAATGCCCATGATACATCCGGGCGCCGATGTGTGTGATCTCTCTGAACTAGGAAATG 1114
Qy      121 LYSLEULYANPPOGILYATGYRGIN1LEMETYGLULYSGLYLYARGSESERGLU 140
Db      1115 AAGTTGAAHAAACCCCGAAGATACAGATCATGAGAGAAAGGCAAGAGCTCATGAG 1174
Qy      141 HIS1LEGLNARG1LEAPATGAPVALSERGLYTHRLEUARGLYHIS1LEPHEPHET 160
Db      1175 CACATCCAGGCGCATTCGCGGACGTPAACCGGAGCACTTAAGAGCAATATATCTCAGG 1234
Qy      161 ASPATGYRGLYTHRILEYSGILNARGILULEULENHIS1LEULEUALATYRGLUGLUTYR 180
Db      1235 GATGATAGCGAACAACAGAGCGGAGACTTACCTCCACATCTCTGGCATATGAGAGTAT 1294
Qy      181 AANPROGLUVALGLYTYRYSARGSPLEUSERHIS1LEALALEUPHELEULEUTYR 200
Db      1295 AACCCGAGAGTGGCTACTGCGAGGACCTGAGCCACATCCGCCCTTGTTCTCTCTAT 1354
Qy      201 LEUPROGLUGLUASPALAPHETPALALEUVALGLULEULEUALASERGLUARGHISER 220
Db      1355 CTTCCTGAGAGAGATGATCTTGTGGACCTGTGTCACTGTGCTGTGCGCATGAGAGGCACTCC 1414
Qy      221 LEUNGILYHNEHISERPROAENGILYGLYTHRVALIGLGLYLEUGLUAAPGLINGILU 240
Db      1415 CTGCGAGGATTTTCACACCCCAATATGGCGGACCGTCTCAAGGGCTTCAAGACCAACAGAG 1474
Qy      241 HISVALVALAATHRSERGINPOLYSTHMETGLYHISGLINAPLYLYASPLEUCYS 260
Db      1475 CATGTGTATACCAAGTCAACACCAAGACATGGGGCATCG----- 1516
Qy      261 GLYGLNYSERPROLEULYCYSEULEIARG1LEULEI1LEAPGLY1LESERLEUGLY 280
Db      1517 -----ATCTCTCTGGG 1528

```

```

QY      281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluGlnAlaLeuMetProIle 300
      1529 CTCACCTCGGCTGGGAGCGTGTATCTGGTGAAGGAGCAAGCGCTTGAATCCGATA 1588
QY      301 ThrArgIleAlaPheLeuValGlnGlnIleValArgLeuThrIleSerArgCysGlyPro 320
      1589 ACAAGATCGCTTTAAAGTTGACGAGAAAGCGCTCCAGAAAGAGCTCAGGTGTGCCCG 1648
QY      321 TrpAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspGluAspThrValLeu 340
      1649 TGGGACGTTTTTGGCAACGGTTCGTGTATACCTGGGCCAGGAGATAGAGCATGTGGCTC 1708
QY      341 LysHleLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
      1709 AAGCATCTTAAGGCTCTTATGAGAACTTACAGAAAGAGGAGGAGCTTGCACACCCCA 1768
QY      361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
      1769 GCCAAACCCGAGCAGGTCGTGGCATCAGGCTGTGGCTTCAAGTGGCGGAGAG 1828
QY      381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProAlaPro 400
      1829 ACCCTCTGCAAGGGGAGCAGGAGCGCTCCAGGCCACAGCCGGTTCGCGCGGCC 1888
QY      401 IleTrpSerAlaSerProProAlaArgAlaProArgSerSerThrProCysProGlyValAla 420
      1889 ATTGGTCAAGCTTCCCGGCAAGGCACTCTGTTCTTCCACACCTGTCTGTGGGGGCT 1948
QY      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
      1949 GTCCGGGAGAGACACTACCTGTGGGACCTCAGGGTGTGCCAGCGCGCTGCTCAG 2008
QY      441 GlyGlyProGlnGlySerTrpArgPheLeuGlnITTPAsnSerMetProArgLeuProThr 460
      2009 GGAAGGACCTCAGGGTTCCTGGAGATTCCTGAGTGAACCTCCATGCCGCCCTCCCAAG 2068
QY      461 AspLeuAspValGluGlyProTrpPheArgHleTyrAspPheArgGlnSerCysTrpVal 480
      2069 GACCTTGAGAGTAAAGGAGCCCTTGTGTCCGCTATTAATTCAGACAGAGCTCTGGATC 2128
QY      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnIleProAlaGln 500
      2129 CGTGCCATATCCAGAGAGGACAGCTGGCCCTGTGGAGGCTGAACCTTGGGAG 2188
QY      501 ArgValaIleSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
      2189 CGGGTGAGATCGGCTTCCCTGCAACAGCAGCATGATTCGACCAAGGAGCAACCCCTTCAG 2248
QY      521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHleAla 540
      2249 GCTAGGAGAGCAAGCGGTGTCTCCACCTCAGGAGCTGTGCTGTGGCTTCCACTTG 2308
QY      541 GluSerSerGlnPheProGlyPhe 549
      2309 GAAAGTTCTCAGTTCCCTCAGGCTTC 2335
Db
RESULT 12
AX775949
LOCUS      AX775949      2647 bp      mRNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 219 from Patent WO03048202.
ACCESSION  AX775949
VERSION     AX775949.1 GI:32693667
KEYWORDS
ORGANISM   Homo sapiens (human)
SOURCE     Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1
            Matsuda, A. and Muramatsu, S.
            NF-KB activating gene
            Patent: WO 03048202-A 219 12-JUN-2003;
            Asahi Kasei Kabushiki Kaisha (JP)

```

```

FEATURES
    source                Location/Qualifiers
                        1..2647
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        755..2338
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAE11674.1"
                        /db_xref="GI:32693668"
                        /translation="MDVAVAGSMVAQEREDILIMKYEKHRAGLBEDKPKPRSTYN
                        NVDHGIVHETELPLTAAEAQIIRREISRSKSWMDMGDEKYSRGLIDRAKGM
                        PMNIRGPMWSVLNTEEMKLNKPRYOIMKEKSKSESHIQRIDRVSQTLKHLIFR
                        DRYGRKRELHILAYERYNEVAVCYRDLSHIALFLILYLEDPAFMLVOLLASER
                        HSLQGFSPNGGTVOGLDQDQSHVAVTSQPTKHOISGLTLRLMDVLYVEBOALM
                        PRTTAAFKQQRRLRTKTSRCGPMAFCNRFVUTYMARDEDTYKHLRASMKLTGRQD
                        LQPPAKPEQSSASRPVPASRGKTLCKDRAPGPPAPRFPRTMSAPPPAPNSST
                        PCPGAVREDIYVPVQGVSPALAQGGQGSWRPLQWNSMRPLTDLDEGPWPAHY
                        DFGSCWVALISQBDOLAPCMQAEHPAERVRSFAFAPSTDSOGTFFRARDQPCAPT
                        SGPCLCGLHLESSQPPPGF"
ORIGIN
Alignment Scores:
Pred. No.:      1,066-97
Score:          2827.00
Percent Similarity: 95.63%
Best Local Similarity: 95.63%
Query Match:    94.77%
DB:             Gaps:          1
US-10-071-838-2 (1-549) x AX775949 (1-2647)
QY      1 MetAspValValGluValAlaGlySerTrpTPAlaGlnGluArgGluAspIleIleMet 20
      755 ATGAGCTGTGTAGAGTCCCGGACAGTTGTGGCCACAGACGAGAGACATCATTATG 814
QY      21 LysTyrGluLysGlyYHIaArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
      815 AATATGAAAGAGGACACCGAGCTGGCTGCCAGAGACAAAGGCGCTTAAGCTTTTGA 874
QY      41 SerTyrAsnAsnAsnValAspHleLeuGlyIleValHleGlnThrGluLeuProPheLeu 60
      875 AGCTACAAACAAACACTCATCTTGGGAGATGTATCAGAAACGAGAGCTGCTCTCAG 934
QY      61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerIleTrpValaAsp 80
      935 ACTGCGCGGAGAGCGAAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGGTGAT 994
QY      81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
      995 ATCTGGAGAGATCGGAGAAATACAAAGACACAGAAAGCTCATGATCGAGCTTACAG 1054
QY      101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlnMet 120
      1055 GGAATGCCATGAACATCCGGGCGCCGATGTGTGTGAGTCTCTCTGAACACTGAGGAAATG 1114
QY      121 LysLeuLysAsnProGlyValArgTyrGlnIleMetLysGlyLysArgSerSerGlu 140
      1115 AAGTTGAAAACCCCGAAGATACAGATCATGAAAGAGAAAGGCAAGGTCATCTGAG 1174
QY      141 HleIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHleIlePhePheArg 160
      1175 CACATCGACGCATCGACCGGAGCGTAAAGCGGAGCATTTAAGAGACATATATCTTCAAG 1234
QY      161 AspArgTyrGlyThrLysGlnArgLysLeuLeuHleIleLeuAlaTyrGluGlnTyr 180
      1235 GATCGATAGGAAACCAAGCAGCGGAGAACTACCTCCATCTCTCGCATATAGAGATAT 1294
QY      181 AsnProGluValGlyTyrCysArgAspLeuSerHleIleAlaLeuPheLeuLeuTyr 200
      1295 AACCCGAGAGTGGCTACTGCAAGGAGCTGACCCACATGCGCCCTGTCTCTCTCAT 1354
QY      201 LeuProGluGluAspAlaPheThrTrpAlaLeuValGlnLeuLeuAlaSerGluArgHleSer 220

```



```
Db      1355 CTTCCTGAGAGAGATGCAATTCCTGSCATGCTGACGCTGCTGCGCAGTGAAGGACATCC 1414
Qy      221  LeuGlnGlyPheHLeuSerProHsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db      1415 CTGCGAGGATTTCAACAGCCCAATGCGCGGACCTCTCCAGGGGCTCCAAAGACCAACAGAG 1474
Qy      241  HisValAlaIleThrSerGlnProLysHrmMetGlyHsnGlnAspLysLysAspLeuGly 260
Db      1475 CATGGTAGCCACGTCACCAACCCAGACCAATGGGCGATCG----- 1516
Qy      261  GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db      1517 -----ATCTCTCTCGGG 1528
Qy      281  LeuThrLeuArgLeuThrAspValTyrLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
Db      1529 CTCACCTTCGCGCTGTGGAGCTGTATCTGTAGAGAGCGAACAAGCGGTGATGCCGATA 1588
Qy      301  ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db      1589 ACAAGATTCGCTTTAAGTTCAAGCAGAACGCTTCAAGAACGTCACAGGTGGCCG 1648
Qy      321  TyrAlaArgPheCysAsnArgPheValAspThrTyrAlaArgAspGluAspThrValLeu 340
Db      1649 TGGCGACGTTTTCGCAACCGGTTCTGTGATACCTGGGCGCAGGATGAGACACTGTGCTC 1708
Qy      341  LysHisLeuAlaArgLysSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db      1709 AAGCATTTAGGGGCTTATGAGAACTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1768
Qy      361  AlaLysProGluGlnGlySerSerLysSerArgProValProLysSerArgGlyLys 380
Db      1769 GCCAAACCCAGAGAGAGGCTGTGGCATCCAGGCTGTGGCTTCAAGTGGCGGAGAG 1828
Qy      381  ThrLeuCysAlaGlyAspArgGlnAlaProProGlyProProAlaArgPheProPro 400
Db      1829 ACCCTTCGACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1888
Qy      401  IleThrSerAlaSerProProAlaArgLysSerSerThrProCysProGlyGlyAla 420
Db      1889 ATTGGTCAAGCTTCCCCGACAGGGACCTCGTTCTTCCACACCTCTGTGGGGGCT 1948
Qy      421  ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db      1949 GTCCGGAGAGACCTAACCTCTGTGGACCTCAAGGGGTGCCCAAGCCCGGCTGTGCTCAG 2008
Qy      441  GlyGlyProGlnGlySerTyrPargPheLeuGlnThrAsnSerMetProArgLeuProThr 460
Db      2009 GAGAGACCTCAAGGCTTCTGTGAGATTCCTGCACTGAGAACCTCAATGCCCGCTCCCAAG 2068
Qy      461  AspLeuAspValGlnGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
Db      2069 GACCTGAGAGCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2128
Qy      481  ArgAlaIleSerGlnGlnLysGlnLysAlaProCysTyrGlnAlaGlnHisProAlaGln 500
Db      2129 CTTGCAATATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2188
Qy      501  ArgValArgSerAlaPheAlaIleProSerThrAspSerAspGlnGlyThrProPheArg 520
Db      2189 CGGGTAGATCGGCTTTGCTGCTCAACCCAGCATGATTCGAGACAGAGGACCCCTTCAGA 2248
Qy      521  AlaArgAspGlnGlnProCysAlaIleProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db      2249 GCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2308
Qy      541  GluSerSerGlnPheProProGlyPhe 549
Db      2309 GAAAGTTCTCACTTCCTCCAGGCTTC 2335
```

RESULT 13
AX833552

```
LOCUS      AX833552                2647 bp    DNA        linear    PAT 15-DEC-2003
DEFINITION Sequence 676 from Patent EPI347046.
ACCESSION  AX833552
VERSION     AX833552.1   GI:39919687
KEYWORDS
SOURCE
  ORGANISM  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
  1. Isegai, T., Sugiyama, T., Otsuki, T., Makamatsu, A., Sato, H., Ishii, S.,
  Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
  Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
  Masuno, Y.
  Full-length cDNA sequences
  Patent: EP 1347046-A 676 24 SEP-2003;
  Research Association for Biotechnology (JP)
FEATURES
  source
    1..2647
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      1.06e-97      Length:      2647
Score:          2827.00      Matches:      525
Percent Similarity: 95.63%      Conservative: 0
Best Local Similarity: 95.63%      Mismatches:   2
Query Match:    94.77%      Indels:       22
DB:              6              Gaps:         1
US-10-071-838-2 (1-549) x AX833552 (1-2647)
Qy      1  MetAspValValGluValAlaGlySerTyrPargLaglnGluArgGluAspIleIleMet 20
Db      755  ATGACGCTGTAGAGAGAGTCCGCGGACGTTGGTGGGCAACAAGCGAGACATCATTAATG 814
Qy      21  LysTyrGlnLysGlyLysArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db      815  AAATACGAAAGAGAGACCGAGCTGGGCTGTCACAGAGACAAAGGGGCTTAAGCCCTTTTCA 874
Qy      41  SerTyrAsnAsnValAspHisLeuGlyIleValHisGlnThrGluLeuProProLeu 60
Db      875  AGCTACAAACAACACGTCGATCATTTGGGATTTGATAGACAGAGCTGCTCTCTG 934
Qy      61  ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerTyrValAsp 80
Db      935  ACTCGGGGAGAGGCGAAGCAAAATTCGCGGGAGATACGCCGAAAGCAAGTGGGTGAT 994
Qy      81  MetLeuGlyAspTyrPargLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db      995  ATGCTGGAGACCTGGAGAAATTCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1054
Qy      101  GlyMetProMetAsnIleArgGlyProMetTyrPargValLeuLeuAsnIleGlnGluMet 120
Db      1055  GGAATGCCCATGAACATCCGGGCGCGAGTGTGCTGCTCTCTGAAACATGAGAGAAATG 1114
Qy      121  LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnLysGlyLysArgSerSerGlu 140
Db      1115  AAGTGAAGAAACCCCGAGAGATACAGATCATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1174
Qy      141  HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db      1175  CACATTCAGGCAATTCACCGGACGTAAGCGGAGCACTTAAGAGAGCATATATCTTCAGG 1234
Qy      161  AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGluTyr 180
Db      1235  GATCGATTACGAGACCAAGCAGCGGAGAACTACCACTCCTCGGCATATGAGAGATAT 1294
Qy      181  AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuTyr 200
Db      1295  AACCCGAGGTGGGCTACTGCAAGAGACCTGAGCCACATCGCCGCTGTCTCTCATAT 1354
```


QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnuArgHisSer 220
 Db 1355 CTTCTGAGAGAGATGATTCCTGGGACCTGGTGCACCTCTGGCCAGTGAAGGACCTCC 1414
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 Db 1415 CTGGAGGGATTTTCAACAGCCCAATAGCGGGACCGTCCAGGGGCTCCAAACCAAGAGG 1474
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysValAspLeuCyS 260
 Db 1475 CATGTGTGATCCACGTCACCAACCAAGACCATGTGGGACATCAG----- 1516
 QY 261 GlyGlnCySerProLeuGlyCySerLeuLeuLeuAlaAspGlyLeuSerLeuGly 280
 Db 1517 -----ATCTCTCTCGGG 1528
 QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
 Db 1529 CTCACCCCTGGCGCTGTGGGACGTGTATCTGTATGAAAGGGAACAGGCGTTGATGCCGATA 1588
 QY 301 ThrArgGlnAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCySerGlyPro 320
 Db 1589 ACAGAATCGCCTTTAAGGTTCAAGCAGAACGCCCTCCACAGACGTCACAGTGTGGCCCG 1648
 QY 321 TrpAlaArgPheCyAspAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 1649 TGGGACCGTTTGTGCAACCGGTTCTGTATACCTGGGCGAGAGAGACCTGTGGCTC 1708
 QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
 Db 1709 AAGCATCTTAGGGCGCTTATGAAGAAACATCAAGAAAGCAGGGGAGCCTGCAACCCCA 1768
 QY 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
 Db 1769 GCCAAACCCGAGCAAGGCTGTGGGATCCAGGCGTGTCCGCTTCAGTGTGGCGAG 1828
 QY 381 ThrLeuCyValArgAlaAspArgGlnAlaProProGlyProProAlaArgPheProPro 400
 Db 1829 ACCCTCTGAGAGGGGACAGGAGGCGCTCCAGGCGCCAGCCCGGTTCCCGGCGCC 1888
 QY 401 IleTProSerAlaSerProProArgAlaProArgSerSerThrProCyAspProGlyLys 420
 Db 1889 ATTGTGTAGCTTCCCGGCGAGGACCTGTTCTTCCACACCTGTCTGTGGGCGT 1948
 QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1949 GTCCGGGAGAGACACCTACCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTCGAG 2008
 QY 441 GlyGlyProGlnGlySerThrArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
 Db 2009 GGAGGACCTCAAGGGTCTCTGGAGATTCTCGAGTGAATCTCAATGCCCCGCTCCCAAG 2068
 QY 461 AspLeuAspValGlyGlyProTrpPheArgHisTyrAspPheArgGlnSerCySerVal 480
 Db 2069 GACCTGGAGCTGAGAGGGCGCTGTGTTCCGCAATTGATTTCAGACAGAGCTGTGGCT 2128
 QY 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCySerGlnAlaGlnHisProAlaGln 500
 Db 2129 CGTGCATATTCACAGAGAGACAGCTGGCGCCCTGCTGGCAGGTCAGAACACCTCGGAG 2188
 QY 501 ArgValaLysSerAlaPheAlaAlaProSerThrArgAspSerArgGlnGlyThrProPheArg 520
 Db 2189 CGGGTGAATCGGCTTTCGCTGACCCAGCACTGATTTCGACACAGGAGCCCTTCCTCA 2248
 QY 521 AlaAspAspGlnGlnProCyValaLysProThrSerGlyProCyLeuGlyLeuHisLeu 540
 Db 2249 GCTAAGGAGAGACAGCCGCTGTCTCCACCTCAGAGGCTTGTGCTGTGGCGCTCCACTTG 2308
 QY 541 GluSerSerGlnPheProProGlyPhe 549
 Db 2309 GAAAGTTCAGATTCCCTCCAGGCTTC 2335

RESULT 14

AK095385

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,

Wakamatsu, A., Hayashi, K., Sato, H., Naga, K., Kimura, K., Makita, H.,

Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,

Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,

Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,

Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,

Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,

Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,

Tanikawa, M., Yamazaki, M., Nimomiya, K., Ishida, T., Yamashita, H.,

Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, S.,

Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,

Yosida, M., Horiuchi, T., Kusano, D., Kanehori, K., Takahashi-Fujii, A.,

Hara, H., Tanase, T., Nomura, Y., Togiyasu, S., Komai, F., Hara, R.,

Takeuchi, K., Ariita, M., Imose, N., Mutsaers, K., Yuuki, H., Oshima, A.,

Saeki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,

Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,

Terasaka, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,

Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,

Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,

Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,

Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,

Ozaki, K., Hiroo, M., Ohnori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,

Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,

Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunuma, K.,

Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, T., Togaishi, T.,

Oyama, M., Hara, H., Watanabe, M., Komatsu, T., Mitushima-Sugano, J.,

Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,

Nagase, T., Nomura, N., Kikuchi, H., Nakagawa, K., Okumura, K.,

Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

2

Tanigami, A., Fujiwara, T., Shidhara, T., Goto, Y., Hiroo, M.,

Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K.,

Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,

Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T.,

Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,

Kikuchi, H., Kanda, K., Wagatsuma, M., Murakami, K., Kanehori, K.,

Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,

Sugano, S., Nagahari, K., Maeuho, Y., Nagai, K. and Isogai, T.

NEO human cDNA sequencing project

Unpublished

3 (bases 1 to 2647)

Isogai, T. and Yamamoto, J.

Direct Submission

Submitted (04-JUN-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7

Kazusa-Kanetani, Kiseazu, Chiba 292-0812, Japan

(E-mail: genomics@kai.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEO human cDNA sequencing project supported by Ministry of

Economy, Trade and Industry of Japan, cDNA full insert sequencing:

Research Association for Biotechnology (RAB); cDNA library

construction: Helix Research Institute (HRI) (supported by Japan

Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,

HRI, and Biotechnology Center, National Institute of Technology and

Evaluation; clone selection for full insert sequencing: HRI and

RAB; annotation: HRI and RAB.

Location/Qualifiers

1..2647

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CTONG2015330"
/issue_type="tongue, tumor tissue"
/clone_id="CTONG2"
/note="Cloning Vector: pME18SFL3"

Alignment Scores:

Pred. No.: 1 06e-97 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.63% Conservative: 0
Best Local Similarity: 95.63% Mismatches: 2
Query Match: 94.77% Indels: 22
Gaps: 1

US-10-071-838-2 (1-549) x AK095385 (1-2647)

QY 1 MetAapValValGluValAlaGlySerTTPTrpAlaGlnGluArgLysAspIleIleMet 20
Db 755 ATGACGTGTAGAGCTCCGGGAGTTGTGGGCAAGAGCCAGAGCATCATTAAG 814
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 815 AAATACGAAAGGAGACACCGAGCTGGGCTGCCAGAGGCAAGAGGCTTAAGCTTTTGA 874
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
Db 875 AGCTACACACACACGTCGATCATTTGGGAGTTGTATCATGAGACGAGCTGCTCTGTG 934
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTyrValAsp 80
Db 935 ACTGGCGGGAGGAGGAGCAAAATTCGGCGGAGATCAGCCGAAAGAGAGTGGTGGAT 994
QY 81 MetLeuGlyAspTTPGluLysTyrLysSerSerArgLysLeuIleAspAlaGlyTyrLys 100
Db 995 ATGCTGGAGAGCTGGGAGAAATCAAAAGACAGACAGAGCTCATAGATCCAGCTACAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
Db 1055 GGAATGCCCATGAAACATCCGGGCGGAGTGTGTCAATCTCTGAAACCTAGGAAATG 1114
QY 121 LysLeuLysAsnProGlyValArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 1115 AAGCTGAAAAACCCCGGAGAGATACCAATCATCTAGAGAGAGGAGCATCTGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATGACCGGAGCGTAAAGCGGACATTAAGAGACATATATCTCAGG 1234
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db 1235 GATCGATACCGAGACCAAGCAGCGGAGAACTACTCCATCTCTCGGCAATTAAGAGAT 1294
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuTyr 200
Db 1295 AACCCGGAGGTGGCTACTGCTGAGGAGACCTGAGCCACATCCGCCCTTGTCTCTCAT 1354
QY 201 LeuProGluGluAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 1355 CTTCCTGAGAGAGATGATCTCTGGCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyLysThrValGlnGlyLeuGlnAspGlnGln 240
Db 1415 CTGCGAGGATTTTCAAGCCCAAAATGGGGGAGCGCTCCAGAGGCTCTCAAGACCAAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 1475 CATGTGTAGGACGTCACCAACCAAGACCATGGGAGTCTGAG----- 1516
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280

Db 1517 -----ATCTCTCCGGG 1528
QY 281 LeuThrLeuArgLeuThrAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 1529 CTCACCTGCTCCCTCTGTGGAGCTGTATCTGTGTAAGAGGAGGAGGAGGCTTATCCATTA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgCysGlyPro 320
Db 1589 ACAAGATATGCTTTTAAAGTTTCAGCAGAGAGCCCTTCAAGAGAGTCCAGGTGTGCCG 1648
QY 321 ThrAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspGluAspThrValLeu 340
Db 1649 TGGGACAGTTTTCACACCGTTCCTTATACCTGGGCGAGGAGTATGAGCACTGTGCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1709 AAGCATCTTAGGCTCTTATAGAAAGAACTAAAGAAAGAGAGGAGGAGCTGCACACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1769 GCCAAACCGAGCAAGAGGTGTGTCGACATCCAGGCTGTGCGCTTCACTGTGGGAGAG 1828
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGCAAGGGGAGCAGGAGGAGCCCTCCAGGAGCCACAGCCGCTTCCGCGAGCC 1888
QY 401 IleTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1889 ATTTGGTACGCTTCCCGGACAGGAGCAGGAGCCTGTTCTTCCACACCTGTCTGTGGGCT 1948
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCCGGAGAAAGACCTTACCTGTGGGACATCAGGATGTCCAGCCCGGCTGTGCTCAG 2008
QY 441 GlyGlyProGlnGlySerTTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
Db 2009 GAGGAGCCTCAGGGTCTCTGAGATTCCTGACAGTGAATCCATGCTCCCGCTTCCAGG 2068
QY 461 AspLeuAspValGlnGlyProTTPPheArgHisTyrAspPheArgGlnSerCysTTPVal 480
Db 2069 GACCTGAGCGTAAGAGGAGCCCTGTGTCCGCTATATGATTTCAGACAGAGCTCTGGGCT 2128
QY 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTTPGlnAlaGlnHisProAlaGln 500
Db 2129 CGTCCATATCCAGAGAGGACCAAGCTGAGCCCTGCTGTGAGGCTTAAACCTGTGGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 2189 CGGGTGAATCCGGCTTTCGTGACCCAGACCTGAATTCGAGACAGAGGACCCCTTCAGA 2248
QY 521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
Db 2249 GCTAGAGAGCAAGACCGCTGTGCTCCACCTCAGGAGGCTTGTGAGGCTTCACTTGA 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 2309 GAAAGTCTCAAGTTCCTCCAGGCTTC 2335

RESULT 15
LOCUS BC058890
DEFINITION Homo sapiens cDNA clone MGC:64921 IMAGE:5744726, complete cds.
ACCESSION BC058890
VERSION BC058890.1 GI:37590617
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,


```
QY      281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGluAlaLeuMetProIle 300
Db      925 CTACCTGCGCCCTCTGAGACGTGTATCTGTAGAGGAGGACAGGCGTTTCATGCTGATA 984
QY      301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db      985 ACMAA-ATCGCTTTTAAGGTTTCAGCAGAGAGCCCTCAGAGAGAGCTCAGGTGTGCCCCG 1043
QY      321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 1043
Db      1044 TGGGCAAGTTTTTGGAAACCGGTTTGGTATGCTGTGGCCAGGGATGATGACACTGTCTC 340
QY      341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGluAspLeuProProPro 360
Db      1104 AAGCATCTTAGGCGCTTATAGAACTAACAGAGAGAGGAGGAGCCTGCACCCCA 1163
QY      361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db      1164 GCCAAACCCGAGCAAGGCGTGTGGCATCCAGGCTGTGCCAGCTTCAAGGGCGGAG 1223
QY      381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db      1224 ACCCTGTGCAAGGGGAGCAAGGAGGCGCTCCAGGCGCCACAGCCCGTTCCCGCGGCC 1283
QY      401 IleTrpSerAlaSerProProAlaArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db      1284 ATTTGGTCACTTCCCGGACAGGCGCACTGTTCTTCCACACCTGTCTCTGGTGGGCT 1343
QY      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db      1344 GTCCGGGAGAGACACTTACCTGTGGCATCTAGGCTGTGCCAGCCCGGCTGTGCTCAG 1403
QY      441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db      1404 GGAGGAGCTCAAGGCTTCTGGAGATTCTCTGAGTGAACCTCATGCCCCCGCTCCCAACG 1463
QY      461 AspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
Db      1464 GACCTGACAGTGGGAGACCTTGTGTTCCGCGTTATGATTCAGACAGAGCTGTGGTCTC 1523
QY      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGln 500
Db      1524 CGTCCATATCCAGAGAGACAGCGGCGCACTGTGGCAGGCTGAAACCTGTGCGAG 1583
QY      501 ArgValArgSerAlaPhe----- 506
Db      1584 CCGGTGAGATCGGCTTTCACTGCACTGAGCCACAAGCTGGGCATGACTTCCCGGCGCTG 1643
QY      507 -AlaAlaProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGluGlnPro 526
Db      1644 CAGTGGCGCCAGCACTGATTCGACAGAGGCACTCCCTTCAGAGCTAGGAGCAAGCA 1703
QY      526 OCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeuGluSerSerGlnPhePro 546
Db      1704 GTGTGCTCCCACTCAAGGCGCTTGTGCTGTGCGGCTCCACTTGGAAAGTTCTCAGTTC 1763
QY      546 OProGlyPhe 549
Db      1764 TCCAGGCTTC 1773
```

Search completed: February 4, 2005, 10:14:09
Job time : 6882 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 05:10:43 ; Search time 753 Seconds
(without alignments)
3827.267 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MOVVEAGSWAQRERDIIM.....TSGPCICGHLHESQPPPGF 549

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+gpn.model -DEV=xlh
-O=/cgn2.1/USPTO_pool/US10071838/runac_03022005_071217_179/app_query.faeta_1.711
-DB=N Genesegq_23Sep04 -QFMT=faetap -SUFFIX=ing -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=-1 -END=-1 -MATRIX=biomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10071838@cgn2.1_1.470@runac_03022005_071217_179 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -MAPN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N Genesegq_23Sep04:*
1: genesegq19808:*
2: genesegq19908:*
3: genesegq20008:*
4: genesegq2001as:*
5: genesegq2001bs:*
6: genesegq2002as:*
7: genesegq2002bs:*
8: genesegq2003as:*
9: genesegq2003bs:*
10: genesegq2003cs:*
11: genesegq2003ds:*
12: genesegq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	6	ABSS4706
2	2983	100.0	1964	10	ADCC37382
3	2978	99.8	1993	3	AAAI5001
4	2975	99.7	1964	10	ADCC37384
5	2963	99.3	2072	5	ABX71403
6	2907.5	97.5	2146	6	ABSS4707

7	2906.5	97.4	2304	10	ADCC37380	Adc37380 Nuclear f
8	2898.5	97.2	2304	10	ADCC37378	Adc37378 Nuclear f
9	2850	95.5	2084	6	ABZ11810	Abz11810 Human pol
10	2850	95.5	2084	12	ADMA44328	Adma44328 Novel hum
11	2835	95.0	2647	10	ADCC37388	Adc37388 Nuclear f
12	2827	94.8	2072	6	ABZ11809	Abz11809 Human pol
13	2827	94.8	2072	12	ADMA44327	Adma44327 Novel hum
14	2827	94.8	2647	10	ADCC37386	Adc37386 Nuclear f
15	2827	94.8	2647	11	ADMO1991	Admo1991 Human CDN
16	2784	93.3	1752	6	ABST8751	Abst8751 DNA encod
17	2773	93.0	1862	6	ABSS4708	Abss4708 CDNA enco
18	2191.5	73.5	7856	5	AA685531	AA685531 DNA encod
19	2191.5	73.5	7856	5	AA688213	AA688213 DNA encod
20	2191.5	73.5	7878	6	AB164684	Ab164684 Stomach c
21	2135.5	71.6	8435	4	AA159633	AA159633 Human pol
22	2135.5	71.6	8435	4	AA159634	AA159634 Human pol
23	2133.5	71.6	8435	10	ADCC31883	Adc31883 Human nov
24	2133.5	71.5	8320	2	AA688212	AA688212 DNA encod
25	2133.5	71.5	8201	2	AA1212170	AA1212170 pJG4-5-CD
26	2133.5	71.5	8201	10	AD685053	Ad685053 Farnesyl
27	2133.5	71.5	8284	5	AA685535	AA685535 DNA encod
28	2133.5	71.5	8408	4	AA157848	AA157848 Human pol
29	2133.5	71.5	8420	4	AA157847	AA157847 Human pol
30	1795.5	60.2	3396	5	AA685534	AA685534 DNA encod
31	1795.5	60.2	3396	5	AA688214	AA688214 DNA encod
32	1732	58.1	2838	5	AA687029	AA687029 DNA encod
33	1732	58.1	2838	5	AA684009	AA684009 DNA encod
34	1732	58.1	2838	5	AA687522	AA687522 DNA encod
35	1533.5	51.4	2424	5	AA687523	AA687523 DNA encod
36	1507.5	50.5	9805	4	AA636324	AA636324 Human car
37	1507.5	50.5	9805	4	AA105169	AA105169 Human rep
38	1507.5	50.5	9805	4	AB198056	Ab198056 Human tes
39	1507.5	50.5	9805	10	ADE47018	Ad47018 Human car
40	1500.5	50.3	9404	4	AA636326	AA636326 Human car
41	1500.5	50.3	9404	4	AA105171	AA105171 Human rep
42	1500.5	50.3	9404	4	AB198058	Ab198058 Human tes
43	1500.5	50.3	9404	10	ADE47020	Ad47020 Human car
44	1469	49.2	9804	4	AA636329	AA636329 Human car
45	1469	49.2	9804	4	AA105173	AA105173 Human rep

ALIGNMENTS

RESULT 1	ABSS4706	strandard; CDNA; 1964 BP.
ID	ABSS4706	strandard; CDNA; 1964 BP.
AC	ABSS4706;	
XX		
DT	28-NOV-2002	(first entry)
XX		
DE	CDNA encoding human PRC17 protein.	
XX		
KW	Human; PRC17; prostate cancer; ovarian cancer; lung cancer;	
KW	breast cancer; cytoskeletal; chromosome 17q11-12; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1650
FT		/*tag= a
FT		/product= "Human PRC17 protein"
XX		
PN	WO200262958-A2.	
XX		
PD	15-AUG-2002.	
XX		
PF	08-FEB-2002; 2002MO-US003457.	
XX		
PR	08-FEB-2001; 2001US-0267615P.	
XX		
XX	(TUL-)	TULARIK INC.
XX		

XX	RESULT 2
XX	ADCC37382
ID	ADCC37382 standard; DNA; 1964 BP.
XX	
XX	ADCC37382;
XX	
DT	18-DEC-2003 (first entry)
DE	
XX	Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 215.
XX	
KW	Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW	cancer; infectious disease; bone disease; AIDS;
KW	neurodegenerative disease; ischemic disorder; Antiinflammatory;
KM	immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
XX	Neuroprotective; Neotropic; Cardiant; Gene therapy; human; gene; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO2003048202-A2.
XX	
PD	12-JUN-2003.
XX	
PE	03-DEC-2002; 2002WO-JP012644.
XX	
PR	03-DEC-2001; 2001JP-00368692.
PR	05-DEC-2001; 2001US-0335829P.
PR	03-OCT-2002; 2002JP-00291302.
PR	04-OCT-2002; 2002US-0415769P.
XX	
PA	(ASAH) ASAMI KASEI KK.
XX	
PI	Matsuda A, Muramatsu S;
XX	
DR	WPI; 2003-505282/47.
XX	P-PSDB; ADCC37383.
PT	
PT	New purified protein that activates nuclear factor kappa B (NF-kappaB) ,
PT	useful for treating inflammation, autoimmune diseases, cancer,
PT	infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT	ischemic disorders.
XX	
PS	Claim 4; SEQ ID NO 215; 938bp; English.
XX	
CC	The present invention relates to novel proteins and their coding
CC	sequences (ADCC37168-ADCC37455), which activate nuclear factor kappa B (NF-
CC	kappaB). The proteins and their coding sequences are useful for treating
CC	a disease associated with NF-kappaB activation, such as inflammation,
CC	autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC	neurodegenerative diseases, or ischemic disorders.
XX	
SO	Sequence 1964 BP; 472 A; 553 C; 561 G; 378 T; 0 U; 0 Other;
	Alignment Scores:
	Pred. No.: 1,42e-128 Length: 1964
	Score: 2983.00 Matches: 549
	Percent Similarity: 100.00% Conservative: 0
	Best Local Similarity: 100.00% Mismatches: 0
	Query Match: 100.00% Indels: 0
	DB: 10 Gaps: 0
US	US-10-071-838-2 (1-549) x ADCC37382 (1-1964)
OY	
1	1 MetAaPvAlValGluValAlaGlySerTPPTAlaGlnGluuArgGluuAspIleIleMet 20
Dh	1 ATGGACGGTGTGAAGGTCGCGGCACTGTGTGGCCAAAGAGGAGGACATATTATG 60
OY	21 LysTYrGluTYrGlyVHISaRgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Dh	61 AAATACGAAAGAGGACACCGAGCTGGCTGCCAGAGGACAAAGGGGCTTAAGCTTTTGA 120
OY	41 SerTYrAaenAaenValAspHISrLeuGlyIleValHISrGluThrGluLeuProPheLeu 60

Db	121	AGCTAACAAACAACAGCTGATCATTTGGGGATTGACATGAGACGGAGCTGCTCTCTG	180
Qy	61	Thrlaarglualalyvglnlleargrglu!eserarglyserlytrpvalasp	80
Db	181	ACTCGCGGGGCGGAGCAAAATTCGCGGGAGATCAACCGMAAGCAAGCGGGGAT	240
Qy	61	Metleuglyasptpglulvetylysersearglybleu!laasphargla!alylys	100
Db	241	ATGCTGGAGACTGGAGAAATATACAAAGACAGAAAGCTCATGATGAGCGTACAG	300
Qy	101	GlymeCPromecaenillearglyPromeCTrpserValleuLeuasnilleglumec	120
Db	301	GGATTCGCCATGAAACATCCGGGGCCCATGtGtCAgTCTCTCTGACATTAAGAAATG	360
Qy	121	Lyaleu!ybaAnProgl!yargTyrgnillewelysgulvrglylyargserseglu	140
Db	361	AAgTTGAAAAACCCCGGAAGATACAGATCATGAAGGAGAAAGGCGMAAGGTCATCTGAG	420
Qy	141	Hist!leGlnargl!leasphargasvalserclYthre!u!aglyvhis!llephentarg	160
Db	421	CACATTCAGCCCATCGAACCGGGAGGTAAAGCGGACATTAAAGAACATATATCTTCAGG	480
Qy	161	AspargtTyrglYthrlYsgln!argGlu!eu!eu!his!leleu!eul!atYrglugluTy	180
Db	481	GATGATACGGAAACCAACAGACGGGGAACCTACCAATCTCTCGCATTAAGAGATAT	540
Qy	181	AsnProgl!uvalGlyTyrcYsargaspleuSerhis!le!al!aleu!phe!eu!eTy	200
Db	541	AAccCGAGGtGGGCTACTGACGGAACTTGAGCCACATCGCGCTTGTCTCTCTCAT	600
Qy	201	LeuProgl!uglu!asphalaphetTPlaleuValGln!eu!eul!aSerGlu!arghisSer	220
Db	601	CTTCTCGAGGAGATGACATCTTGCGCATCTGTGACGCTGCGCGACAGAGGACATCC	660
Qy	221	LeuglnglyPhehisSerProasnGlyGlyThrValGlnGlyLeuGlu!aspslnglnglu	240
Db	661	CTGACAGGATTTCCACAGCCCAAAATGGCGGAGCCGTCCAGGGGCTCCAAACCAAGAG	720
Qy	241	H!sValValal!athrSerGlnPro!yathrmerGlyhis!sglnaspy!aly!aspleu!cys	260
Db	721	CATGTGGTAGCACTGCACAAACCAAGCCATGGGGCATCAGGACAAAGAAATCATAT	780
Qy	261	GlyGlnCysSerPro!euglyCysleu!lleargl!eleu!lleaspgly!lleSer!eugly	280
Db	781	GGGAGAGTCTCCCGTTAGGCTGCTCATCCGGAATTAATTAAGGGAATCTCTCGGG	840
Qy	281	Leuthre!u!arg!e!u!TPaspy!aly!re!u!valGlnGlyGln!al!e!u!merPro!le	300
Db	841	CTCACCCCTGGCGCTGTGGAGCTGATCTGTAAGAGCGAACAGCGCTGATGCCATTA	900
Qy	301	Thzarg!leal!aphe!ysValGlnGln!y!sarg!e!u!Thr!Ser!Arg!Cys!gl!Pro	320
Db	901	ACAGAATTCGGCTTTAAGTTCAACAGAGGCGCTCACGAAGAGCTCCAGAGGTGGCCG	960
Qy	321	TTP!laargPheCysasnargPheVal!aspThr!TP!laargaspGlu!asptThr!Val!eu	340
Db	961	TGGCAGAGTTTTCCAACCGGTTGTTGATCTCTGGGCGACAGAGATGAAGACACTGTGCTC	1020
Qy	341	Lyhis!le!u!arg!al!aser!Me!ly!le!u!Thr!arg!ys!Gln!y!asp!leu!Pro!Pro	360
Db	1021	AAAGCATCTTGGGCGCTCTATGAAAGAAATTAACAAAGAAAGCGGGGACCTGCCACCCCA	1080
Qy	361	Al!ay!sProgl!uglnglyser!Ser!al!aser!argPro!al!aser!Arg!yGly!ys	380
Db	1081	GCCAAACCCGAGCAAGGCTGTGGCATTCAGGCGCTGTGCCGCTTCAAGTGGCGGAAG	1140
Qy	381	Thrl!euCys!ys!gl!y!aspa!rg!ln!al!asPro!Progl!y!Pro!Pro!al!arg!Phe!Pro!arg!Pro	400
Db	1141	ACCTCTGCAAGGGGAGACAGGCAAGCCCTTCAGGCGCACAGCCCGGTTCCCGGGGCC	1200
Qy	401	lle!tr!ser!al!aser!Pro!Pro!arg!al!Pro!arg!Ser!er!Thr!Pro!Cys!Progl!yGly!al!	420
Db	1201	ATTGTGTAGCTTCCCGCACCGGCACTGTGTTCTTCACACCCCTGTGCTGTGTGGGCT	1260

QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1261 GTCCGGGAAACACTAACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCTGTGCTCAG 1320
 QY 441 G1G1ProGlnIySerTPrpAhpLeuGlnITPrpAsnSerMetProArgLeuProThr 460
 Db 1321 GGAGGACCTCAGGGTTCCTGGAGATTCTCGCAGGTGAATCTCATGCCCGCTCCCAAG 1380
 QY 461 AspLeuAspValGlnGlyProTTrpPheArgHsITyTrpAspPheArgGlnSerCysTrpVal 480
 Db 1381 GACCTCGACGTAGAGGGCCCTTGTTCCGCATTAGATTTCAGACAGCGTGTGGGGTC 1440
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHsIProAlaGlu 500
 Db 1441 CATTCCATTATCCACAGAGACCAAGCTGGCCCTCTGTGGCAGGGCTGAAACCTTCGGAG 1500
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerArgGlnGlyThrProPheArg 520
 Db 1501 CGGGTGAAGATGGGCTTTGGCTGTCAACCACTAATTCGACCAAGGGCACCCCTTCAGA 1560
 QY 521 AlaArgAspGlnGlnProCysAlaAlaProThrSerGlyProCysLeuGlyLeuHsIleu 540
 Db 1561 GCTAGGGAGCAACAGCCGTGTGTCTCCACCTCAAGGGCTTGTCCCTCGGGCTTCCACTTG 1620
 QY 541 GluSerSerGlnPheProProGlyPhe 549
 Db 1621 GAAAGTTCTCAGTTCCTCCAGGGCTTC 1647

RESULT 3
AAA15001
ID AAA15001 standard; cDNA, 1993 BP.
XX

AC	AAA15001;
XX	
DT	21-AUG-2000 (first entry)
XX	
DE	cDNA encoding a human proliferation and apoptosis related protein.
XX	

KM nullum, proliferacion and apoptosis related protein, PROAP, psoriasis,
 KM cell proliferative disorder, immunological disorder, hepatitis,
 KM reproductive disorder, arteriosclerosis, cirrhosis, melanoma,
 KM cancer, acquired immune deficiency syndrome, AIDS, allergy, anaemia,
 KM asthma, diabetes mellitus, osteoarthritis, endometriosis,
 KM uterine fibroid, menstrual cycle, ss.
 XX Homo sapiens.

PT for diagnosis, treatment and prevention of cell proliferative,
PT immunological and reproductive disorders.
XX

The present sequence encodes a human proliferation and apoptosis related protein (PROAP). The polypeptides and polynucleotides can be used for the diagnosis, treatment and prevention of cell proliferative, immunological and reproductive disorders. Disorders associated with decreased expression or activity of include arteriosclerosis, cirrhosis, hepatitis, psoriasis, melanoma, lymphoma and cancers of the breast, brain and prostate, acquired immune deficiency syndrome (AIDS), allergies, anaemia, CC asthma, diabetes mellitus, osteoarthritis, endometriosis, uterine CC fibroids and disruptions of the menstrual cycle. Antibodies against PROAP CC can be used in diagnosis of disorders characterized by PROAP e.g. in ELISAs CC (enzyme linked immunosorbent assays) and the polynucleotides may be used CC to detect and quantify gene expression in biopsied tissues. These CC techniques can also be used to monitor regulation of PROAP levels during CC therapeutic intervention

Sequence 1993 BP; 478 A; 564 C; 570 G; 381 T; 0 U; 0 Other;

Alignment Scores:	
Pred. No.:	2.45e-178
Score:	2978.00
Percent Similarity:	99.83%
Best Local Similarity:	99.82%
Query Match:	99.83%
DB:	1
Length:	1993
Matches:	548
Conservative:	0
Mismatches:	1
Indels:	0

US-10-071-838-2 (1-549) x AAA15001 (1-1993)

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspLleIleMet 20

Db 42 ATGGAGCGTGTAGAGTGTGGCGGCGAGTTGGTGGGCGACAGAGCGAAGGACATCAATATG 107

QY 21 LysTYrGluLysGlyNH₂ArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40

Db 102 AAATTCGAAAAGGACACCGAGCTGGGGCTGCCAGAGAGCAAGGGGCTTAAGCCTTTTCGA 167

QY 41 SerTYrAsnAsnAsnValAspHisIleLeuGlyIleValHisGluThnGluLeuProProLeu 60

Db 162 AACTTCACAACAACGTCGATCATTTGGGGATTTGTACATGAGACGAGCCTCCTCTCTG 221

QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

Db 222 ACTGGCGGGAGGGCGAAGCAAAATTCGGCGGGAGATCAGCCGAAGAGCAAGCTGGTGAT 281

QY 81 MetLeuGlyAspTrpGluLysTYrLysSerSerArgLysLeuIleAspArgAlaTYrLys 100

Db 282 ATGGCTGGGAGCTGGGGAATTCAAAGACAGCAAAAGCTCATAGATTCAGCGCTACAG 341

QY 101 GlyMetProMetSerIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120

Db 342 GGATTCGCCCAATGAACATCCGGGGCCCGATGTGTGTCAGTCTCTGAACTAGAGGAATG 401

QY 121 LysLeuLysAsnProGlyArgTYrGlnIleMetLysGlnLysGlyLysArgSerSerGlu 140

Db 402 AAGTTGAAAAACCCCGAGATTCACAGATCTGAAGAGAAAGGACAGAGCTCATCTGAG 461

QY 141 HisIleGluArgIleAspArgArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160

Db 462 CACATCCACGCGATGACCCGGGAGCGTAAGCGGGACATTAAAGAAAGCATATATTCTTCAG 521

QY 161 AspArgTYrGlyThrLysGlnArgGluLeuLeuHisGlyLeuLeuAlaTYrGluGluTYr 180

Db 522 GATCGATACGGAACCAACACACGGGAATCTCTCCACATCTCTCGCATATAGAGGATAT 581

QY 181 AsnProGluValGlyTYrCysArgAspLeuSerHisIleAlaIlePheLeuLeuTYr 200

Db 582 AACCCGGAGGTGGGTACTGCGAGGAACTTAGCCACATGCGCCCTTGTCTCTCTCAT 641

QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220


```

Db      642  CTCTCGAGAGAGATGCAATTCGGGCACTGGTGCAGCTGCTGGCCAGTGAAGGCACTCC
Oy      221  LeuGlnGlyPheHisSerProbaGlnGlyThrValGlnGlyLeuGlnAspGlnGln
Db      702  CTGCAGGATTCATTCAGCCCAATGCGCGGACCGCTCCAGGGGCTCCAGAACAGAGAG
Oy      241  HisValValAlaThrSerGlnProIleThrMetGlnHisGlnAspGlyLeuAspLeu
Db      762  CATGTGTAGCCACCTCAACACCCAGAACCAATGGGCACTGAGGACAGAAATCTAGT
Oy      261  GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu
Db      822  GGGCAGGTGTCCTCCGCTTAGGCTGCTCATCCGGATTTGATGAGGGATCTCTCCGG
Oy      281  LeuThrLeuArgLeuTrpAspValIleuValGlnGlyGlnAlaLeuMetProIle
Db      882  CTCACCTGCGCTGTGGGACGTGTATCTGTAGAAAGCGAACAGCGCTGTATGCCGATA
Oy      301  ThrArgIleAlaPheLeuValGlnGlnIleValAspGlnThrLeuThrSerArgCysGlyPro
Db      942  ACAAGAAATCGCTTTAAGGTTCAAGCAAGAGCGCTCAAGAACGTCACAGGTGAGCCCG
Oy      321  TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGlnAspThrValIleu
Db      1002  TGGGACGTTTTCGCAACCGCTTCCTGATACCTGGGCCAGGGATGAGGACACTGTGCTC
Oy      341  LysHisLeuArgAlaSerMetLysIleuThrArgLysGlnGlyAspLeuProPro
Db      1062  AAGCATCTTAGGGCTCTATGAAAGAACTAACAAAGAGAGAGGAGGAGCTCCACCCCA
Oy      361  AlaIysProGlnGlnGlySerSerIleAspArgProValProIleSerArgIleGlyLys
Db      1122  GCCAAACCCGAGCAAGGCTGCTGGCATTCAGGCGCTGTCGGCTTCACGTGCGGAGAG
Oy      381  ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro
Db      1182  ACCCTCTGCAAGGGGAGCAAGGAGCGCTCCAGGCCACCAAGCCCGTCCGGGGGCC
Oy      401  IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla
Db      1242  ATTTGTCAGCTTCCCGCCAGCGGACCTCTTCCACACCTGTCTGTGGGGGCT
Oy      421  ValArgGlnAspThrTrpProValGlnGlnGlnValProSerProAlaLeuAlaGln
Db      1302  GTCCGGAGAGACACTACCTGTGTGGCACTCAGGGGTGTCAGGCCGCTGACTCAG
Oy      441  GlyGlyProGlnGlySerTrpArgPheLeuGlnIleProAsnSerMetProArgLeuProThr
Db      1362  GAGAGACCTCAGGGTTCCTGGAGATTCCTGCAGTGGAACTCAATGCCCGCTCCCAAG
Oy      461  AspLeuAspValGlnGlyProTrpPheArgHisIleAspPheArgGlnSerCysTrpVal
Db      1422  GACCTGAGAGTGAAGGCGCTTGGCTCCGCACTTATGATTTGACAGAGCTGCTGGGTC
Oy      481  ArgAlaIleSerGlnGlnAspGlnIleuAlaProCysTrpGlnAlaGlnHisProIleGln
Db      1482  CGTGCATATCCAGAGAGAGACAGCTGGGCCCTCTGTCAGAGCAACCTCCGAGG
Oy      501  ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg
Db      1542  CGGGAGATGCGGCTTTCCTGCTCACCAGCACTGATTCGACAGGAGGACCCCTTCAGA
Oy      521  AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu
Db      1602  GCTAAGGAGCAACACCGCTGTCTCCACCTCAGGGGCTTGGCTGCGGCTCCACACTTG
Oy      541  GluSerSerGlnPheProProGlyPhe 549
Db      1662  GAAAGTTCAGATTCCTCCACAGGCTTC 1688

```

```

ID      ADc37384 standard; DNA; 1964 BP.
XX
AC      ADc37384;
XX
DT      18-DEC-2003 (first entry)
XX
DE      Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 217.
XX
KW      Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW      cancer; infectious disease; bone disease; AIDS;
KW      neurodegenerative disease; ischemic disorder; Antiinflammatory;
KW      Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW      Neuroprotective; Nootropic; Cardiac; Gene therapy; human; gene; de.
XX
OS      Homo sapiens.
XX
FN      WO2003048202-A2.
XX
PD      12-JUN-2003.
XX
PF      03-DEC-2002; 2002WO-JP012644.
XX
PR      03-DEC-2001; 2001JP-00368692.
XX      05-DEC-2001; 2001US-0335829P.
XX      03-OCT-2002; 2002JP-00291302.
XX      04-OCT-2002; 2002US-0415769P.
XX
PA      (ASAH) ASAH1 KASEI KK.
XX
PI      Matseuda A, Muramatsu S;
XX
DR      MPI; 2003-505282/47.
XX      P-PSDB; ADc37385.
XX
PT      New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT      useful for treating inflammation, autoimmune diseases, cancers,
PT      infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT      ischemic disorders.
XX
PS      Claim 4; SEQ ID NO 217; 938bp; English.
XX
SQ      Sequence 1964 BP; 473 A; 552 C; 561 G; 378 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,31e-128 Length: 1964
Score: 2975.00 Matches: 548
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.73% Indels: 0
DB: 10 Gaps: 0
US-10-071-838-2 (1-549) x ADc37384 (1-1964)
Oy      1  MetAspValValGlnValAlaGlySerTrpTrpAlaGlnGlnIleuArgIleLeuMet 20
Db      1  ATGACGTCGTGAGAGTTCGGGCGAGTTCGTGGGACAGAGCCAGAGACATCATTTTG 60
Oy      21  LysTyrGlnLysGlyHisIleArgAlaGlyLeuProGlnAspLysGlyProLysProPheArg 40
Db      61  AAATAGGAAAGGAGCACCGAGCTGGCTGCAGAGCAAGGGGCTTAAGCTTTTGA 120
Oy      41  SerTyrAsnAspAsnValAspHisIleuGlyIleValHisGlnThrGlnLeuProLeu 60
Db      121  AGCTACACACACAGCTGATTCATTTGGGAGATTTGACATGAGACGAGACTGCTCTCTG 180
Oy      61  ThrAlaArgGlnAlaLysGlnIleArgArgGlnIleSerArgLysSerLysTrpValAsp 80

```

```

Db      181  ACTGCGCGGAGGAGCAACCAATTGCGGGAGATCGCCGAAAGACCAAGTGGTGGAT 240
Qy      81  MetLeuGIYAAPTPTGluLeuTyrTyrLeuSerSerArgTyrLeuLeuAspArgAlaTyrTyr 100
Db      241  ATGCTGGAGAGCTGGGAGAAATACAAAGCAGAGAAAGCTCATATGATGAGGATCAAG 300
Qy      101  GluMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGluMet 120
Db      301  GGAATGCCCATGAACTATCCGGGGCCCATGTGGTCTCTCTGAAACATTAATGAGAAATG 360
Qy      121  LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLeuGlyLysArgSerGlu 140
Db      361  AAGTTGAAAAACCCCGAAGATACAGATCATAGAGAGAGGAGAGGATCATGAG 420
Qy      141  HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db      421  CACATCCAGCGCATTCGACCGGAGCTTAAGCGGAGCATTAAGAGACATATATCTTCAGG 480
Qy      161  AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
Db      481  GATCGATACGAAACCAAGCAGCGGGAACTACTCCATCTCTCGCATATGAGAGATAT 540
Qy      181  AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTyr 200
Db      541  AACCCGAGAGTGGGCTACTGACGAGCACTGAGCCACATGCGCGCTTCTCTCTAT 600
Qy      201  LeuProGluGluAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db      601  CTTCCTGAGAGAGATGCAATTCGGGCACTGGTGCAGCTGCTGGCATGAGAGCACTCC 660
Qy      221  LeuGlnGlyPheHisSerProAsnGlyGlyTyrValGlnGlyLeuGlnAspGlnGln 240
Db      661  CTGCGAGGATTTCAAGCCCAATGCGGAGCCGTCACAGGAGGCTCCAAAGCCAAAGAG 720
Qy      241  HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db      721  CATGTGGTACCAAGTCAACAACCCAGACCATGGGAGCATCAGACAAAGAAATCATAT 780
Qy      261  GlyLysCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLleSerLeuGly 280
Db      781  GGGCAGTGTCCCGTTAGGCTGCTCATCCGATATGATGACGAGCATCTCTCCGG 840
Qy      281  LeuThrLeuArgLeuTyrPheValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db      841  CTCACCTCGCCCTGTGGAGCTGTATCTGTAGAAGCAGAACGCTTATATGCCATTA 900
Qy      301  ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db      901  ACAAGATCGCTTTTAAGTTTCAGAGAACGCTCTCAGAGACGTCACAGGTGTGGCCG 960
Qy      321  TyrAlaArgPheCysAsnArgPheValAspThrTyrAlaAspAspGluAspThrValLeu 340
Db      961  TGGGCACTTTTTCACACCGGTTCTTATATCTGGCCAGAGATGAGAGACCTGTCTC 1020
Qy      341  LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db      1021  AAGCATCTTAAAGGCTTATAGAACTAACAGAAAGCAGGGGAGCCTGCAACCCCA 1080
Qy      361  AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db      1081  GCCAAACCCGAGCAAGGGTGTCTGCGCATTCAGGCTGTGCGGCTTCAGTGGCGGAG 1140
Qy      381  ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db      1141  ACCCTGCAAGGGGAGCAGGAGCCCTCTCAGGCCCAACCCGCTTCCGCGGCC 1200
Qy      401  IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db      1201  ATTGTGATAGTTCCCGCAGCGGACCTGTCTTCCACACCCCTGCTGTGGGGCT 1260
Qy      421  ValArgLysAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

```

```

Db      1261  GTCCGGAGACACTACCTGTGGGCACTAGAGGTGTGCCAGCCCGGCTGCTCAG 1320
Qy      441  GlyValProGlnGlySerTyrPheLeuGlnTyrPheAsnMetProArgLeuProThr 460
Db      1321  GAGAGACTTACAGGGTTCGAGAGATTCGACGTGAACTCATGCCGCCCTCCCAAG 1380
Qy      461  AspLeuAspValGlnGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
Db      1381  GACCTGAGAGTGAAGGCGCTTGTGTCGCTATATGATTTTCAAGACAGAGCTGTGGGTC 1440
Qy      481  ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrPheGlnAlaGlnIleProAlaGln 500
Db      1441  CGTGCATATCCCAAGAGACCAAGTGGCCCTGCTGGCAGGCTGAACCCCTGGGAG 1500
Qy      501  ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db      1501  CGGCTGAGATCGGCTTGTGTCAGCCAGCACTGATTCAGACAGGAGACCCCTTCAGA 1560
Qy      521  AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db      1561  GCTAGAGACCAAGCCGTGTGTCTCCACTCAGAGGCTTGCCTGCGGCTCCACTTG 1620
Qy      541  GluSerSerGlnPheProGlyPhe 549
Db      1621  GAAAGTCTCAGTCCCTCCAGGCTTC 1647

```

RESULT 5
ABX71403
ID ABX71403 standard; cDNA; 2072 BP.
XX
AC ABX71403;
XX
DT 14-APR-2003 (first entry)
XX
XX Human cell cycle-associated cDNA from clone DKFZphc3_35p22.
XX
XX Human; gene; gene therapy; vaccine; disease treatment; detection; ss.
XX
XX Homo sapiens.
XX
XX PN M0200112659-A2.
XX
XX 22-FEB-2001.
XX
XX 18-AUG-2000; 2000MO-IB001496.
XX
XX 18-AUG-1999; 99US-0149499P.
XX
XX 28-SEP-1999; 99US-0156503P.
XX
XX (GHEU-) GERMAN HUMAN GENOME PROJECT.
XX
XX Wiemann B.
XX
XX MPI; 2001-327840/34.
XX
XX P-PSDB; ABUS3234.
XX
XX Nucleic acids having the sequences of clones isolated from libraries of
XX different human tissues, useful in recombinant DNA methodologies.
XX
XX Claim 1; Page 867; 1095pp; English.
XX
XX This invention describes novel polynucleotides and polypeptides isolated
XX from human cDNA libraries which can be used for gene therapy or in
XX vaccines. The polynucleotides of the invention and antibodies encoded by
XX them may be used in the prevention, diagnosis and treatment of diseases
XX associated with inappropriate polypeptide expression. The products of the
XX invention may also be used to identify modulators of expression and
XX activity and to down regulate expression and activity. The antibodies of
XX the invention may also be used as diagnostic agents for detecting the
XX presence of polypeptides in samples. This sequence encodes a polypeptide
XX described in the disclosure of the invention
XX
XX Sequence 2072 BP; 510 A; 579 C; 586 G; 397 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,24e-127 Length: 2072
 Score: 2963.00 Matches: 545
 Percent Similarity: 99.64% Conservative: 2
 Best Local Similarity: 99.27% Mismatches: 2
 Query Match: 99.33% Indels: 0
 DB: 5 Gaps: 0

US-10-071-838-2 (1-549) x ABX71403 (1-2072)

```

QY 1 MetAspValValGluValAlaGlySerTPTPAlaGlnGluArgGluAspIleIleMet 20
DB 99 ATGAGAGTGTAGAGTCCGCGGCACTGTGGGCAAGAGGAGAGGACATCATTAAG 158
QY 21 LysTyrGluLysGlyValArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 159 AAATACGAAAGGAGGACCGAGCTGGGCTGCAGAGGACAAAGGGGCTTAAGCTTTTCCG 218
QY 41 SerTyrAspAspAspValAspHisLeuGlyIleValHisGluThrGluLeuProProLeu 60
DB 219 AGCTACAAACAACACGTCATCTGGGATTTGTACATGAGACGAGAGCTGCTCTCTG 278
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTyrValAsp 80
DB 279 ACTGCGCGGAGGCGGAGGAAATTCGCGGAGATCAGCCGAAAGGCAAGTGGGTGAT 338
QY 81 MetLeuGlyAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 339 ATGCTGGAGAGCTGGGAGAAATACAAAGACGAGAAAGCTCATATCGAGGCTACAG 398
QY 101 GlyMetProMetAspIleArgGlyProMetTyrSerValLeuLeuAspIleGluGluMet 120
DB 399 GGAATGCCATGAAACATCCGGGCGCGAGTGTGAGTCTCTCTGAACTGAGAAAG 458
QY 121 LysLeuLysAspProGlyValArgTyrGlnIleMetLysGluLysGlyLysSerSerGlu 140
DB 459 AAGTTGAAAGAACCCCGGAGATACAGATCATAGAGAGAGGCAAGAGTATCTGAG 518
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 519 CACATCAGCGCATTCAGCGGAGCTAAGCGGAGCATTAAGAAAGATATATCTTCAGG 578
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
DB 579 GATCGATACGGAACCAAGCAGCGGAACTACTCCATCTCTGCGATATGAGAGTAC 638
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
DB 639 AACCGGAGGTGGGCTATCTGAGGAGACTGAGGCACTGCGGCTTTGCTCTCTAT 638
QY 201 LeuProGluGluAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 699 CTTCCTGAGAGAGATCATCTCTGGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 758
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
DB 759 CTGAGAGGATTTACAGCCCAATAGCGGAGCGTCCAGGAGCTCCAAAGCAACAGAG 818
QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
DB 819 CATGTGTAGCCACGTCACACCAACCAAGACATGGGGCATAGAGCAAGAAAGATCTAT 878
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB 879 GGGCAGATGTTCCCGTTAGGCTGCTCATCCGATATATGATGACGGAGATCTCTCGGG 938
QY 281 LeuThrLeuArgLeuThrAspValTyrLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
DB 939 CTCACCTCGGCTGTGGAGCGTGTATCTGTGTAAGAGGCAACAGCGCTGATGCGGAT 998
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
  
```

```

DB 999 ACNAGATGCGCTTTAAGGTTGACAGAAAGCGCTTCACGAGACGTCAGGTGTGCGCG 1058
QY 321 ThrAlaArgPheCysAspAspArgPheValAspThrTPAlaArgAspGluAspThrValLeu 340
DB 1059 TGGGCGCTTTTGTGCAACCGGTTGTTGATACCTGGGCGCAAGGATAGAGACATGTGCTC 1118
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 1119 AAGCATCTTAAGGGCCCTTATGAAGAACTTAACAAGAAAGAGGGGACCTGCGCACCCCA 1178
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
DB 1179 GCCAAACCCGAGCAAGGGTGTGTGGCATCAGGCTGTGCGGCTTCACTGTGCGGAGAG 1238
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1239 ACCCTTGAAGGGGAGGAGGAGGCGCTTCCAGGCGCACAGCCGGTTCGCGGAGCC 1298
QY 401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyLysAla 420
DB 1299 ATTGGTCACTTCCCGGCAAGGACCTGCTTCCACACCTGTCTGTGGGGGCT 1358
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1359 GTCCGGGAGAGACACTAACCTGTGGGCACTCAGGGTGTGCCAGCCCGGCGCTCAG 1418
QY 441 GlyGlyProGlnGlySerThrArgPheLeuGlnThrAspSerMetProArgLeuProThr 460
DB 1419 GAGAGCTTCAGGGTCTGTGAGATTCTGCAGTGAACTCCATGCGCGCTCCAGAG 1478
QY 461 AspLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
DB 1479 GACCTGAGCTAGAGGCGCTTGTGTTCCGCAATTATGATTTCAGACAGAGCTGTGGTCT 1538
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnHisProAlaGlu 500
DB 1539 CGGTCATATCCAGAGAGGACAGCTGGCGCCCTGTGGCAGGCTGAACACCTGTGGAG 1598
QY 501 ArgValArgSerAlaPheHisAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1599 CGGGTAGATCGGCTTGTGCTGACACAGACATGATTCGACAGGAGCACCCCTTCAGA 1658
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
DB 1659 GCTAGGAGAGAACAGAGATGTCTCCACTCAGAGGCTTGTGCTGCGGCTTCCACTTGG 1718
QY 541 GluSerSerGlnPheProGlyPhe 549
DB 1719 GAAAGTTCTCAGTTCCCTCCAGGCTTC 1745
  
```

RESULT 6
 ABS54707
 ID ABS54707 standard; cDNA; 2146 BP.
 XX
 AC ABS54707;
 XX
 DT 28-NOV-2002 (first entry)
 XX
 XX cDNA encoding human PRCL7 protein splice variant 1.
 DE
 XX Human; PRCL7; prostate cancer; ovarian cancer; lung cancer;
 XX breast cancer; cytostatic; chromosome 17q11-12; gene; 88; splice variant.
 XX
 OS Homo sapiens.
 OS
 XX
 XX
 XX
 FT key Location/Qualifiers
 FT CDS 1..1832
 FT /tag= a
 FT /product= "human PRCL7 protein splice variant 1"
 FT /transl_except= (pos:421..422, aa:Asp)
 FT /note= "This codon has an apparent 1 nucleotide deletion
 FT which alters the reading frame"
 XX

PN M0200262958-A2.
 XX 15-AUG-2002.
 PD 08-FEB-2002; 2002WO-US003457.
 XX 08-FEB-2001; 2001US-0267615P.
 XX (TULAR-) TULARIK INC.
 PA Li J, Powers S, Xiang P, Peng Y;
 XX WPI, 2002-706902/76.
 DR P-BEDB; ABG70737.
 XX
 PT Novel isolated PRC17 polypeptide useful diagnostically or prognostically
 PT to detect diseases or conditions associated with altered PRC17 activity
 XX or expression relative to normal, for example cancer.
 PS Claim 24; Page 63-64; 78pp; English.
 CC The present invention relates to a new PRC17 polypeptide. The invention
 CC is useful for detecting cancer cells (such as prostate tissue, breast
 CC tissue, lung tissue, ovarian tissue) in a biological sample. The
 CC invention is further useful for monitoring the efficacy of a therapeutic
 CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
 CC breast cancer). The molecules of the invention are useful diagnostically
 CC or prognostically to detect diseases or conditions associated with
 CC altered PRC17 activity or expression relative to normal, for example
 CC cancer. The present nucleic acid sequence encodes the human PRC17 protein
 CC splice variant 1 of the invention. The human PRC17 gene is located on
 XX chromosome 17q11-12
 XX
 SQ Sequence 2146 BP; 494 A; 619 C; 616 G; 417 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,54e-125	Length:	2146
Score:	2907.50	Matches:	547
Percent Similarity:	89.67%	Conservative:	0
Best Local Similarity:	89.67%	Mismatches:	2
Query Match:	97.47%	Indels:	62
	6	Gaps:	1

US-10-071-838-2 (1-549) x ABSS4707 (1-2146)

Qy	1	MetAapValValGluValAlaGlySerTrpTrpAlaGlnGluArgGluAspIleIleMet	20
Db	1	ATGGACGTGTAGAGGTGCGCGGCGAGTGGTGGGCAACAGAGGAGAGGACATCATATG	60
Qy	21	LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg	40
Db	61	AAATACGAAAGGAGACACCGAGCTGGCTGCCAGAGGACAAAGGGCTTAAAGCTTTTGA	120
Qy	41	SerTyrPhePhePheValAlaPheHisLeuGlyIleValHis	53
Db	121	AGCTAACAAACAACGTCATCTTTGGGGAATTGACAGTCCCTCGCTGGAGTCA	180
Qy	53	-----	53
Db	181	GCCCCACAGAGAGCCCTTCT	240
Qy	53	-----	53
Db	241	CTGGAAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	300
Qy	54	-----	54
Db	301	GCCCCGTGACAGTCTCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	360
Qy	60	LeuThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerIleTrpVal	79
Db	361	CTGACTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT	419

Qy	80	AspMetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyr	99
Db	420	GATATGCTGGAGAGCTGGAGAAATACAAAGAGAGGAAAGCTCATGATGAGCGTAC	479
Qy	100	LysGlyMetProMetLeuIleArgGlyProMetTrpSerValLeuLeuAsnIleGlnGlu	119
Db	480	AAGGAATGCCCAATGACATCCCGGGCCGATGTGTGCTCTCTCTCTCTCTCTCTCTCTCT	539
Qy	120	MetLysLeuLysAsnProGlyArgTyrGlnIleMetLysGlyLysGlyLysArgSerSer	139
Db	540	ATGAAATTGAAAAACCCCGGAGATACAGATCATGAAAGAGAGGAGGAGGAGGAGGAGG	599
Qy	140	GluHisIleGlnArgIleAspArgPheValSerGlyThrLeuArgLysHisIlePhePhe	159
Db	600	GAGACATCCAGGAGCATGACCGGAGCTTAAGCGGAGCATTAAGAGAGCATATATCTTC	659
Qy	160	ArgAspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGlu	179
Db	660	AGGATCGATACGAGAACCAACAGCGGGAATCTTCCATCTCTGGCATATGAGAG	719
Qy	180	TyrAsnProGluValAlaTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeu	199
Db	720	TATTAACCGGAGGTGGCTACTGAGGACCTAGGCAATGCGCGCTTGTCTCTCTC	779
Qy	200	TyrLeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis	219
Db	780	TATCTTCTGAGGAGATGATCTGCGCATCTGCGCATCTGCGCATCTGCGCATCTGCGCAT	839
Qy	220	SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLysGlnAspGlnGln	239
Db	840	TCCCTGAGGAGATTTTCAAGCCCAATGGCGGAGCGTCCAGGCGCTTCAAGCAACAG	899
Qy	240	GluHisValValAlaHisSerGlnProLysThrMetGlnHisGlnAspLysLysAspLeu	259
Db	900	GAGCATGTGTGAGCGATGACCAACCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	959
Qy	260	CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeu	279
Db	960	TGTGGCAATGTTCCTCGGTAGGCTGCTCATCGGATATTAATTAACGAGATCTCTCTC	1019
Qy	280	GlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnIleAlaLeuMetPro	299
Db	1020	GAGCTCACCTCGCGCTGTGGAGATGATCTGTGTGTAAGGAGGAGGAGGAGGAGGAGG	1079
Qy	300	IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGly	319
Db	1080	ATAACAAATCGCTTTTAAGTTTCAAGAGGCGCTCACAAACATTCAGAGGTGGC	1139
Qy	320	ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal	339
Db	1140	CGGTGGCAGGTTTGGCAACCGTTTCGTGATACCTGAGCCAGAGATGAGGACACTGTG	1199
Qy	340	LeuLysHisLeuArgLysSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro	359
Db	1200	CTCAGAGCTTTAGGCTCTTATGAAACCTAACAAAGAGGAGGAGGAGGAGGAGGAGGAG	1259
Qy	360	ProAlaLysProGlnGlnGlySerSerAlaSerArgProAlaProAlaSerArgGlyGly	379
Db	1260	CCAGCCAAACCCGAGCAAGGAGTGTGCGCATCCAGGCTGTGCGCTTCAAGTGGCGG	1319
Qy	380	LysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg	399
Db	1320	AAGACCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1379
Qy	400	ProIleTrpSerAlaSerProArgAlaProArgSerSerThrProCysArgProGlyGly	419
Db	1380	CCCATTTGTGTAGCTTCCCGGACCGGACCTCGTTCTTCCACACCTGTCTGCGTGGG	1439
Qy	420	AlaValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAla	439
Db	1440	GCTGTCCGAGAAACATCCTCTGTGGACTCAAGGTGTGCCAGCGCGCTGTGCT	1499
Qy	440	GlnGlyValProGlnLysSerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuPro	459

DB 1500 CAGGGAGAGCTCAGGGTTCTGGAGATTCTGCACTGAACTCCATGCCCCGCTCCCA 1559
 QY 460 ThrAspLeuAspValGluGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrp 479
 DB 1560 ACCGACCTGAGCGTAGAGGCGCTTGCTGCCCATTAATGATTCAGACAGAGCTGCTGG 1619
 QY 480 ValArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAla 499
 DB 1620 GTCCGTCCTATATCCAGAGAGACAGCTGGCCCCCTGCTGGAGGCTGAACACCTGTGG 1679
 QY 500 GluArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPhe 519
 DB 1680 GAGCGCGTAGATCGGCTTTCCTGTCACCCAGACACTGATTCGACCAAGGCGACCCCTTC 1739
 QY 520 ArgAlaArgAspGlnGluProCysAlaProThrSerGlyProCysLeuGlyCysGlyLeuHis 539
 DB 1740 AGAGCTAGGAGGAGCAAGCCGTGTGCTCCACCTCAGGCGCTTGCTGCGGCTTCAC 1799
 QY 540 LeuGluSerSerGlnPheProGlyPhe 549
 DB 1800 TTGGAAAGTTCTCAGTTCCCTCCAGGCTTC 1829

RESULT 7

ADC37380

ID ADC37380 standard; DNA; 2304 BP.

AC ADC37380;

DT 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 213.

KM Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KM cancer; infectious disease; bone disease; AIDS;
 KM neurodegenerative disease; ischemic disorder; Anti-inflammatory;
 KM immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
 KM Neuroprotective; Neotropic; Cardiant; Gene therapy; human; gene; ds.

OS Homo sapiens.

PN WO2003048202-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.

PR 03-DEC-2001; 2001JP-0036892.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

PA (ASAH) ASAMI KASEI KK.

PI Maeda A, Muramatsu S;

DR WPI; 2003-505282/47.

DR P-PSDB; ADC37381.

PT New purified protein that activates nuclear factor kappa B (NF-kappaB),

PT useful for treating inflammation, autoimmune diseases, cancers,

PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or

PT ischemic disorders.

PS Claim 4; SEQ ID NO 213; 938bp; English.

CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37380-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischemic disorders.

SQ Sequence 2304 BP; 569 A; 610 C; 636 G; 489 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,4e-125 Length: 2304
 Score: 2906.50 Matches: 540
 Percent Similarity: 98.36% Conservative: 0
 Best Local Similarity: 98.36% Mismatches: 2
 Query Match: 97.44% Indels: 8
 DB: 10 Gaps: 1

US-10-071-838-2 (1-549) x ADC37380 (1-2304)

QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluArgLeuAspIleLeuMet 20
 DB 362 ATGACGTCGTGAGAGTCCCGGCGAGTTGTGGCGACAAGCGAGAGGACATCTTATG 421
 QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 DB 422 AATATGAAAAGGACACCGAGCTGGCTGCCAGAGACAAAGGGGCTTAAGCTTTTCGA 481
 QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
 DB 482 AGCTACAAACAAACAGTCGATCATTTGGAGATTGATGACAGACGAGCTGCTCTCTG 541
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
 DB 542 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCAACCGAAAGCAAGCGGTGAT 601
 QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
 DB 602 ATGCTGGAGAGATCGGAGAAATACAAAGACAGCAAAAGCTCATGATGACGTACAG 661
 QY 101 GlyMetProMetAsnIleArgLysProMetTrpSerValLeuLeuAsnIleGluGluMet 120
 DB 662 GGAATGCCATGAAACATCCGGGCGCGATGTGTGATCTCTTAACACTAGGAAATG 721
 QY 121 LysLeuLysAsnProGlyValGlyGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 DB 722 AAGTTGAAAACCCCGGAGAAATACCAATCATGAAAGAGAGGCGCAAGGTCATCTGAG 781
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 DB 782 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTAAAGAAACATATATCTTCAAG 841
 QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluLysTrp 180
 DB 842 GATCATGAGGAACCAACAGCGGAGACTATCCACATCTCTGTCATATAGAGATAT 901
 QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaLeuPheLeuLeuTrp 200
 DB 902 AACCGGAGTGGCTTACTGACGGACCTGAGCCACATCGCCCTTGTCTCTCTAT 961
 QY 201 LeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlyLysArgHisSer 220
 DB 962 CTTCCTGAGAGAGATGATTTCTGGCCTGGGACGTGCTGCGACAGAGAGCACTCC 1021
 QY 221 LeuGlnGlyPheHisSerTrpAsnGlyLysThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 DB 1022 CTGCAAGGATTTCAACAGCCCAAAATGGCGGACCGGTCCAGGGGCTCCAAAGCAAGAGG 1081
 QY 241 HisValValAlaLeuSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 DB 1082 CATGTGTAGCCACATCACACCCCAAGACATGGGACATCAGACAAAGAAATCTATCT 1141
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
 DB 1142 GGGAGATGTTCCCGCTTAGGCTGCTCATCCGATATGATGAGCGGATCTCTCGGG 1201
 QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGluGlyGluGlnAlaLeuMetProIle 300
 DB 1202 CTCACCTCGCGCTGTGGAGCGTGTATCTGTGTGAAGGAGCAAGCGGTGTATGTCGCAATA 1261
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

```

Db      1262 ACAAGATGCGCTTTAAGTTGAGCAGAAAGCGCTTCCAGAGACCTCCAGGTGTGGCCG 1321
Qy      321 TTPAlaArgPheCysAsnArgPheValAspThrTPAlaArgAspGluAspThrValLeu 340
Db      1322 TGGGACGTTTTTGGCAACCGGTTCCGTGATACCTGGGCAAGGATGAGACACGTGCTC 1381
Qy      341 LysHLeuValArgAlaSerMetLysLysLeuThrArgLysGlnLysAspLeuProPro 360
Db      1382 AAGCATCTTAGGGCTCTATGAAAGAACTTAACAAGAAAGCAGGGGAGCTGCCACCCCA 1441
Qy      361 AlaLysProGluGlnGlnLysSerSerAlaSerArgProValProAlaSerArgLysLys 380
Db      1442 GCCAAACCCGAGCAGAGGCTGTCGCGACCTCCAGGCTTGCGGCTTCCAGTGGCGGAAAG 1501
Qy      381 ThrLeuGlyArgLysArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db      1502 ACCCTTCGACAGGGGAGCAGGCGAGCCCTCCAGGCTTCCAGGCTTCCAGGCTTCCAGG 1540
Qy      401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysArgProGlyVal 420
Db      1541 TT-TGGTCAAGCTTCCCGCCACCGGACCTCGTTCTTCCACACCTCGTCTGAGGAGCT 1599
Qy      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db      1600 GTCCGGGAGACACCTAACCTGTGGGCACTCAGGGTGTGCCAGCCGCGCTGGCTCAG 1659
Qy      441 G1G1YProGlnLysSerTyrPheLeuGlnTyrPaenSerMetProArgLeuProThr 460
Db      1660 GAGAGACCTCAGGGTCTCGAGATTCCTGCAAGTGAATCACTCAATGCCCGCTCCCAAG 1719
Qy      461 AspLeuAspValGlnGlyProTyrPheArgHsrTyrAspPheArgGlnSerCysTyrVal 480
Db      1720 GACCTGAGCGTGAAGGCGCTTGTGTCGCGCATTTGATTTCAACAGAGACCTGTGGGCT 1779
Qy      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnLysProAlaGln 500
Db      1780 CGTGGCATATCCCAAGAGACCAAGTGGCCCTCGTGGCAGGCTGAACCTCTGGGAG 1839
Qy      501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db      1840 CGGGTGAATCGGCTTGTGCTGCACCCAGCATGATTTCCAGACAGGAGCACCCTTCAGA 1899
Qy      521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHisLeu 540
Db      1900 GCTAGGAGCAACAGCCGTTGTCTCCACCTCAGGGCTTGTGCTGCGGCTTCCACTTG 1959
Qy      541 GluSerSerGlnPheProProGlyPhe 549
Db      1960 GAAAGTTCTCAGTTCCTCCAGGCTTC 1986

```

RESULT 8

AD37378 standard; DNA; 2304 BP.

AD37378;

18-DEC-2003 (first entry)

Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 211.

Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease; cancer; infectious disease; bone disease; AIDS; neurodegenerative disease; ischemic disorder; Anti-inflammatory; Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV; Neutroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.

Homo sapiens.

WO2003048202-A2.

12-JUN-2003.

```

PF      03-DEC-2002; 2002WO-JP012644.
XX
PR      03-DEC-2001; 2001JP-00368692.
PR      05-DEC-2001; 2001US-0358829P.
PR      03-OCT-2002; 2002JP-00291302.
PR      04-OCT-2002; 2002US-0415769P.
XX
PA      (ASAH ) ASAH KASEI KK.
XX
PI      Matcoda A, Muramatsu S;
XX
DR      WPI; 2003-505282/47.
DR      P-PSDB; AD37378.
XX
PT      New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT      useful for treating inflammation, autoimmune diseases, cancers,
PT      infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT      ischemic disorders.
XX
PS      Claim 4; SEQ ID NO 211; 938bp; English.
XX
CC      The present invention relates to novel proteins and their coding
CC      sequences (AD37168-AD37455), which activate nuclear factor kappa B (NF-
CC      kappaB). The proteins and their coding sequences are useful for treating
CC      a disease associated with NF-kappaB activation, such as inflammation,
CC      autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC      neurodegenerative diseases, or ischemic disorders.
XX
SQ      Sequence 2304 BP; 570 A; 609 C; 636 G; 489 T; 0 U; 0 Other;

```

Alignment Scores:

```

Pred. No.: 1,26e-124 Length: 2304
Score: 2898.50 Matches: 539
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 3
Query Match: 97.17% Indels: 8
DB: 10 Gaps: 1

```

US-10-071-838-2 (1-549) x AD37378 (1-2304)

```

Qy      1 MetAspValAlaGluValAlaGlySerTyrTPAlaGlnGluAspLysIleMet 20
Db      362 ATGACGCTGTGAGAGTGGCGGCGGCGAGTTGTGTGGGCAAGGCGCAGGACATCATATG 421
Qy      21 LysTyrGlnLysGlyValArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db      422 AATATGAAAGGAGCACCGAGCTGGCTGCAGAGGACAAAGGGGCTTAAGCTTTTCA 481
Qy      41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGlnThrGluLeuProPheLeu 60
Db      482 AGCTACAAACAACAAGCTGAGTATTTGGGATTTGTATGATGACGAGCTGCTCTCTG 541
Qy      61 ThrAlaArgGluAlaLysGlnLysLeuArgGlnLysSerArgLysSerTyrTPValAsp 80
Db      542 ACTGGCGGAGGCGGAGCAAAATTCGCGGAGATCAGCCGAAGAAGCAAGTGGTGAAT 601
Qy      81 MetLeuGlyAspTyrGlnLysTyrLysSerSerArgLysLeuLysAspArgAlaTyrLys 100
Db      602 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAGAGCTCATAGATCGAGCTTCAAG 661
Qy      101 GlyMetProMetAsnLysLeuArgLysProMetTyrPheSerValLeuLeuAsnLysGlnMet 120
Db      662 GGAATGCCATTAACATCCGGGAGCCGATGTGTGAGTCTCTCGAACAACAGAGAAATG 721
Qy      121 LysLeuLysAsnProGlyArgTyrGlnLysLeuLysGlnLysArgSerSerGlu 140
Db      722 AAGTTGAAAAACCCGGAAGATACAGATCATGAAGAGAGAGGCAAGAGCTATCTGAG 781
Qy      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db      782 CATATCAGCGCATGACCGGAGCGTAAGCGGAGCATTAAGAGCAATATATCTTCAG 841
Qy      161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuMetAlaTyrGlnGluLys 180

```

```

Db      842 GATCGATAGGAAACAGACAGCGGGAACTACTCAGATCTCTGGCATATAGAGGATAT 901
Qy      181 AenProGluVal1GlyTyrCyAArgPheLeuSerHis1Ala1AlaLeuPheLeuTyr 200
Db      902 AACCCGAGAGTGGGCTACTGACAGGAGCTGAGCCATGAGCCGCTGTCTCTCTAT 961
Qy      201 LeuProGluGluAspAlaPheTTrpAlaLeuVal1GlnLeuLeu1AspG1uArgHisSer 220
Db      962 CTTCCTGAGAGAGATGATTTCTGGGCACTGGTGCACTCTGCGCATAGAGGACATCTC 1021
Qy      221 LeuGlnGlyPheHisSerProAsnGlyGlyTyrVal1GlnGlyLeuGlnAspGlnGln 240
Db      1022 CTGCAAGGATTTACAGCCCAATAGCCGGGACCGTCCAGGGGCTCCAAACCAAGAG 1081
Qy      241 HisVal1Ala1LeuSerGlnProLysTyrMetGlyHisGlnAspLysLysAspLeu 260
Db      1082 CATGTGTGATGACCAAGCTCACACCAAGACCATGGGGCATCAGGACAAAGAAATCTAT 1141
Qy      261 GlyGlnCySerProLeuGlyCySerLeu1Leu1Leu1Leu1Leu1Leu1Leu1Leu 280
Db      1142 GGGGAGTGTCTCCCGTTAGGCTGCTCATCCGATATGATTTGACGGGATCTCTCTCG 1201
Qy      281 LeuTTrpLeuArgLeuTTrpAspValTyrLeuVal1GlnGlyGlnGln1LeuMetPro 300
Db      1202 CTCACCTGCGCGCTGTGGGACGTGATCTGTGTAAGAGGGAACAGGCGTGTATGCCG 1261
Qy      301 ThrArg1Lea1PheLeuVal1GlnGlnLysArgLeuTTrpLysTTrpSerArgCyG 320
Db      1262 ACMAAATCGCTTTAAGCTTACGACAGAGCCCTTCAACMAACGCTCCAGTGTGGCCG 1321
Qy      321 TrpAlaArgPheCyAsnArgPheVal1AspTTrpAlaArgAspGluAspTrpVal1 340
Db      1322 TGGGACCGTTTGTGCAACGGTTCGTGATACCTGGGCAAGGAGTGAAGACCTGTCTC 1381
Qy      341 LysHisLeuArgAlaSerMetLysLysLeuTTrpArgLysGlnGlyAspLeuProPro 360
Db      1382 AAGCATCTTAGGGCCCTATGAAAGAACTAACAAAGAAACAGGGGACCTTGCAACCCCA 1441
Qy      361 AlaLysProGluGlnGlySerSerAlaSerArgProVal1ProLysArgGlyGlyLys 380
Db      1442 GCCAAACCCGAAACAGGCTGCTGCGCATCAGGCGCTTCCGCGCTTCAAGTGGCGG 1501
Qy      381 ThrLeuCyLysGlyAspArgGln1AlaProProGlyProProAlaArgPheProArg 400
Db      1502 ACCCTGTGAAAGGAGAGAGGAGGCGCTTCCAGAG-----CCA 1540
Qy      401 IleTTrpSerAlaSerProProArgAlaProArgSerSerThrProCyAspGlyGly 420
Db      1541 TT-TGGTCAAGCTTCCCGCCGACAGGCGACCTGCTTCCACACCGCTGCTGGGGCT 1599
Qy      421 ValArgGluAspTrpTTrpProVal1GlyTyrGlnGlyVal1ProSerProAlaLeu 440
Db      1600 GTCGCGGAGAGACACCTACCTGTGGGCACTCAGGGGTGCTCCAGCGCGCTGAG 1659
Qy      441 GlyGlyProGlnGlySerTTrpArgPheLeuGlnTTrpAsnSerMetProArgLeuPro 460
Db      1650 GAGAGACCTCAGAGGTTCTGAGATTCTGCAGTGAACTCCATGCCCCGCTCCCAAG 1719
Qy      461 AspLeuAspVal1GlyTyrProTTrpPheArgHisTyrAspPheArgGlnSerCyTrp 480
Db      1720 GACCTGAGAGATGAGGCGCTTGGTTCGCCCATATGATTTCAGACAGAGCTGCTGG 1779
Qy      481 ArgAla1LeuSerGlnGluAspGlnLeuVal1ProCySerTTrpGlnAlaGlnHisPro 500
Db      1780 CCGTCATATCCAGAGAGAGACAGCTGCGCCCTGCTGCAAGGCTGAAACACCTCGG 1839
Qy      501 ArgVal1ArgSerAlaPheAla1AlaProSerTTrpAsnSerArgGlnGlyTTrpPro 520
Db      1840 CCGGAGTGAATCGGCTTCCGTCGACCAACCATGATTCGAGACCAAGGCGACCCCT 1899
Qy      521 AlaArgAspGluGlnProCyVal1AlaProTTrpSerGlyProCyLeuCyGlyLeuHis 540

```

```

Db      1900 GCTAGGAGCAACACCGCTGTGCTCCACCTCAGGAGGCTTGGCTGCGGCTC 1959
Qy      541 G1uSerGlnPheProProGlyPhe 549
Db      1960 GAAGTTCAGATTCCTCCAGGCTTC 1986

RESULT 9
AB211810
ID AB211810 standard, cDNA; 2084 BP.
XX
XX AB211810;
AC
XX 20-JAN-2003 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 692.
DE
XX
XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cyrostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antarthritic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200270539-A2.
XX
XX 12-SEP-2002.
XX
XX 05-MAR-2002; 2002MO-US005095.
XX
XX 05-MAR-2001; 2001US-00799451.
XX
XX (HVS8-) HVS8Q INC.
XX
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QH, Ren F;
XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
XX Weihsan T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2002-759812/82.
XX
XX P-PSDB; ABP69593.
XX
XX
XX PT New polynucleotides comprising sequences assembled from expressed
XX sequence tags (ESTs), useful for treating cell-proliferative,
XX neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
XX or coagulation disorders.
XX
XX Claim 1; SEQ ID NO 692; 1012pp + Sequence Listing; English.
XX
XX
XX The invention relates to an isolated polynucleotide (1) comprising a
XX nucleotide sequence selected from any of 948 sequences (AB21119-
XX AB212066) or their mature protein coding portion, active domain coding
XX protein or complementary sequences. The polynucleotides are useful for
XX identifying expressed genes or for physical mapping of human genome. The
XX encoded polypeptides (ABP6902-ABP69849) are useful as molecular weight
XX markers, as a food supplement, for generating antibodies, in medical
XX imaging, screening and diagnostic assays and for treating cell-
XX proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
XX or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
XX diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
XX platelet or coagulation disorders, wound, burns, incision, ulcers, liver
XX or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
XX arthritis, etc. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
XX
XX Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,946-122 Length: 2084
XX Score: 2850.00 Matches: 528

```


Percent Similarity: 96.36%
 Best Local Similarity: 96.17%
 Query Match: 95.54%
 DB: 6

Conservative: 1
 Mismatches: 2
 Indels: 18
 Gaps: 1

US-10-071-838-2 (1-549) x AB211810 (1-2084)

QY 1 MetAspValValGluValAlaGlySerTPRTPAAlaGlnGluArgGluAspIleIleMet 20
 Db 163 ATGAGCGTGTATAGGTCCGGGCAATGTGTGGGACCAAGACGAGAGACATCATTTATG 222
 QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 Db 223 AAATACGAAAGGAGACACGAGCTGGCTGCCAGAGGACAAAGGGCCCTTAAGCTTTTGA 282
 QY 41 SerTyrAsnAsnAsnValAspHisIleuGlyIleValHisGluThrGluLeuProPheLeu 60
 Db 283 AGCTACACACACACCTCATCATTTTGGGATTTGATGAGACGAGACCTCCTCCTG 342
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysThrValAsp 80
 Db 343 ACTGGCGGAGGAGCCAGCAATTCGGCGGAGATCACTCCAAAGAGCAAGTGGGTGAT 402
 QY 81 MetLeuGlyAspTyrGluLysThrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
 Db 403 ATGCTGGGAGACTGGGAGAAATACAAAGACAGCAAAAGCTCATAGATCCGACGTACAAG 462
 QY 101 GlyMetProMetAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGluMet 120
 Db 463 GGAATGCCCATGACATCCGGGGCCCGAGTGTGTCACTCTCCCTGAACTGAGGAATG 522
 QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysSerSerGlu 140
 Db 523 AAGTTG----- 528
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 529 AACATCCAGCCCATCGACCGGGAGCTAAGCGGAGCAATTAAGAGCAATATATCTTCAG 588
 QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
 Db 589 GATGATACCGGACCAAGACGCGGAACTCATCTCAATCTCTGCGCAATATGAGAGAT 648
 QY 181 AsnProGluValGlyTyrCysAspAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
 Db 649 AACCCGAGGTGGGCTACTGACAGGACTGAGCAATGCGCCCTTGTCTCTCTAT 708
 QY 201 LeuProGluGluAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 709 CTTCCTGAGGAGATGATGATCTGGGACATGTGCACTGCTGCGCAGTGAAGGCACTCC 768
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 Db 769 CTGCAAGGATTTCAAGCCCAATGCGGAGACCGTCCAGGGGCTCAAGACCAACAGAG 828
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 829 CATGTGTATGCAAGTCAACACCAAGACATGGGAGCATCGAGCAAGAAAGATCTATGT 888
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 889 GGGCAGTGTTCCTCGTTAGGCTGCTCATCCGGAATTTGATGACGGGATCTCTCTGGG 948
 QY 281 LeuThrLeuArgLeuTyrAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
 Db 949 CTCACCTGGCGCTGTGGACCTGTATCTGTAGAGGCGAAGCGGCTGATCCCATTA 1008
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 1009 ACAAGATTCGCTTTAAAGTTTCAAGAGAGCGCTTCAGAGAGACGTCAAGTGTGGCCG 1068
 QY 321 TrrAlaArgPheCysAsnArgPheValAspThrTrrAlaArgAspGluAspThrValLeu 340

Db 1069 TGGCAGCTTTTTCGACCGGTTTCGTTGATACCTGGGCCAGGAGATAGAGACACTGTGCTC 1128
 QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
 Db 1129 AACGATCTTATGGGCTCTATGAAAGAACTTAACAGAAAGCAGGGGACCTCGACACCCCA 1188
 QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
 Db 1189 GCCAAACCCGAGCAGAGGTCGTGGATCCAGGCTGTGGCGGCTTCACTGTGGCGGAAG 1248
 QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1249 ACCCTTCGAAAGGGGACAGGAGGCCCTCCAGGCCCAACAGCCCGGATTCGCGGCCC 1308
 QY 401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyValAla 420
 Db 1309 ATTTGTTCAGCTTCCCGCCACGCGGACCTGTCTTCCACACCTGTCTGTGGGGCT 1368
 QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 1369 GTCCGGAAAGACACTTACCTGTGGGCACTAGAGGTGTCCACGCCCGGCTGTGCTCAG 1428
 QY 441 GlyGlyProGlnGlnGlySerTyrArgPheLeuGlnIleTyrAsnSerMetProArgLeuProThr 460
 Db 1429 GAGGAGCTCAGGGTCTCTGAGATTCCTGAGTGAATCTGAGTGAATCTCAATGCCCGCTCCAAAG 1488
 QY 461 AspLeuAspValGluGlyProTyrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
 Db 1489 GACCTGAGCGTAGAGGGGCCCTGTGTTCCGCAATTAATGATTTCAAGACAGAGCTGTGGTTC 1548
 QY 481 ArgAlaIleSerGlnGlnAspGlnLeuAlaProCysTyrGlnAlaGluHisProAlaGlu 500
 Db 1549 GTGTCAATATCCAGAGGAGACAGCTGGCCCTCTGTGGAGGCTGAACCCCTGGGAG 1608
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 Db 1609 CGGCTGAGATCGGCTTTCGTGCAACCAAGACATGATTCGACCAAGGAGCAACCCCTTCAGA 1668
 QY 521 AlaArgAspGlnGlnProCysValAlaProThrSerGlyProCysLeuCysGlyLeuHisIleu 540
 Db 1669 GCTAGGAGCAACAGCGGTGTCTCTCCACTGACCTGCGCTGTGCGGCTCCACTTG 1728
 QY 541 GlusSerSerGlnPheProGlyPhe 549
 Db 1729 GAAAGTTCATGATTCCTCCAGGCTTC 1755

RESULT 10
 ID ADM44328 standard; cDNA; 2084 BP.
 XX ADM44328;
 AC
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Novel human arginine-rich protein cDNA #692.
 KW ss; gene; human; arginine-rich protein; cancer; inflammation;
 KW genetic disorder.
 XX
 OS Homo sapiens.
 PN US2004053250-A1.
 XX
 PD 18-MAR-2004.
 XX
 PF 21-NOV-2002; 2002US-00302172.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 PR 05-MAR-2002; 2002WO-US005095.
 PR 20-AUG-2002; 2002US-00225251.
 XX
 PA (TANG/) TANG Y T.
 PA (XUEA/) XUE A.

PA (DRMA/) DRMANAC R T.

XX Tang YT, Xue A, Drmanac RT.

XX WPI, 2004-238579/22.

PT New isolated arginine-rich protein-like polynucleotides and polypeptides, useful for diagnosing and/or treating conditions associated with aberrant activity of the arginine-rich polypeptides, such as cancer and inflammation.

PT Inflammation.

PS Disclosure; SEQ ID NO 692; 51bp; English.

XX The invention relates to an isolated polynucleotide. The methods and compositions of the present invention are useful for the diagnosis and/or treatment of diseases or conditions associated with aberrant expression or activity of the arginine-rich protein-like polypeptides, such as cancer and inflammation. They can also be used in forensics, gene mapping, identification of mutations responsible for genetic disorders, and in assessing biodiversity. The present sequence represents a novel human arginine-rich protein cDNA.

XX Sequence 2084 BP; 519 A; 571 C; 581 G; 413 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1 94e-122	Length:	2084
Score:	2850.00	Matches:	528
Percent Similarity:	96.36%	Conservative:	1
Best Local Similarity:	96.17%	Mismatches:	2
Query Match:	95.54%	Indels:	18
DB:	12	Gaps:	1

US-10-071-838-2 (1-549) x ADM44328 (1-2084)

```

QY 1 Metaspvalaigluvalaaglysertrtpalaglnuargluaspjleimet 20
DB 163 ATGAGCTGTAGAGGTCCGGGCACTTGTGGGCAAGAAGGAGACATCATTTATG 222
QY 21 Lvgtrgluluglyhlaarglaaglyleuprogluaspjlygylprolyspoblearg 40
DB 223 AATATCGAAGAGGACCGAGCTGGCTGCCAGAGACAAAGGGCTTTTCA 282
QY 41 Sertraspasnaashvalasphlsleuglylevalhlglnthrgluupproleu 60
DB 283 ACTGACAAACAACATCTCATCTTTGGGATTTGATGACGAGCTGCTCTCTG 342
QY 61 ThrAlaarggluAlaIyaglnleargagluileserarglyserlystrpValaap 80
DB 343 ACTGGCGGAGGCGAAGCAATTCGGCGGAGATCAGCCGAAAGCAAGTGGTGGAT 402
QY 81 Metleuglyaspttrpglulysrlyseraserarglyleuileaspargalalylys 100
DB 403 ATGCTGGAGAGCTGGGAGAAATACAAACACAGCAAGAACTCATATGAGGCTCAAG 462
QY 101 GlymetPrometAsnillearglylpromettrpserValleuileuanilleuglumet 120
DB 463 GGAATGCCCATGACATCCGGGCGCCGATGTGCTCTCTCTGAACTGAGAAAG 522
QY 121 LysleuLysAsnProglYargTrglnIlemetLysglulysglYlysargSerSerGlu 140
DB 523 AGTTG----- 528
QY 141 HlsllleaglNarglleaspaArgaspaValserGlyThleuNarglyhlslllephlearg 160
DB 529 AACATCTCAGCGCATTCACCGGAGCGTAAGCGGAGCACTTAAGAAAGCATATATCTTCAAG 588
QY 161 AaspaArglyGlyThlyserglNarggluLeuLeuHlsllleuLeuAlaTyrgluglylyr 180
DB 589 GATTCATACGAACCAAGACCGGAGAACTACTCAATCTCTTGGCAATGAGAGTAT 648
QY 181 AasProgluValaIyTyrcYaspaArgLeuSerHlsllleAlaIaleupheleuLeuTy 200
DB 649 AACCGGAGAGTGGGCTACTGAGGAGCACTGAGCCATCGCGCTTGTCTCTCTAT 708

```

```

QY 201 LeuProgluGluAspaAlaPheTrpAlaLeuValaGlnLeuLeuAlaSerGluArgHlsSer 220
DB 709 CTTCCTGAGAGAGATGATCTTGGGCACTGGTGCACTGCTGGCCAGTGAAGGCACTTC 768
QY 221 LeuglnglyPheHlsSerProaanglyGlyThrValaGlnglyLeuGlnaaspGlnglnglu 240
DB 769 CTGAGGAGATTCACAGGCCAAATGGCGGAGCGTTCAGAGGGCTCCAAAGCAAGAG 828
QY 241 HlsValaAlaThrSerGlnProlystrMetGlyHlsGlnaaspjlyaspaLeuCy 260
DB 829 CATGGGTGACGACGACCAACCAAGCAAGGAGGAGTCAAGACCAAGAAATCATAT 888
QY 261 GlyGlnCySerProleuGlyCySerleuilearglleuileaspGlylleserleugly 280
DB 889 GGGCAATGTTCCCTTAGGCTGCTCATCCGATTTATGAGGAGTCTCTCGG 948
QY 281 LeuThrLeuArgLeuTrpAspaValTyrlleuValaGlnglyGlnGlnaLeuMetProile 300
DB 949 CTCACCTTCGCGCTGTGGGACGTATCTGTAGAAAGGAGCAAGGCGTGTATGCCGATA 1008
QY 301 ThrArglleAlaPheLysValaGlnGlnlyAspLeuThrLysThrSerArgCySerGlyPro 320
DB 1009 ACAAGATTCGCTTTAAGGTTCAAGCAAGCCCTCAACCAAGACGTCAGGTGGCCG 1068
QY 321 TrpAlaArgPheCyAsnaArgPheValaAspThrTrpAlaArgAspaGluAspThrValleu 340
DB 1069 TGGGACGTTTGTGCAACCGGTTCTGTGATCTGGGCAAGGAGATGAGACATGTGCTC 1128
QY 341 LysHlsleuArgAlaSerMetLysLysLeuThrArglyGlnglyAspLeuProPro 360
DB 1129 AAGCATCTTAGGGCTCTATGAAAGAACTPAACAAGAAAGCAAGGGGAGCTGCACCCCA 1188
QY 361 AlaLysProGlnGlnGlySerSerAlaSerArgProValProAlaSerArglyGlyLys 380
DB 1189 GCCAAACCGACCAAGGCTGCGGATCCAGGCTGTCGCGCTTCAAGTGGCGGAG 1248
QY 381 ThrleuCyLysGlyAspaArglnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1249 ACCCTGTGAAAGGAGACAGGAGGCTCTTCAAGGCCACAGCCGGTCTCCGAGGCC 1308
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCyAspProGlyAla 420
DB 1309 ATTGGTACGTTTCCCGGCAAGGAGCACTGTTTTCACACCCGTGCTGTGGGGCT 1368
QY 421 ValArgGluAspThrTyrlProValaGlyThrglnglyValProSerProAlaLeuAlaGln 440
DB 1369 GTCCGGGAAGACACCTACCTGTGGGCACTCAGGGGTGTGCCAGCCCGGCTGTGCTCAG 1428
QY 441 GlyGlyProGlnGlySerTrpArgPheleuGlnTrpAsnSerMetProArgLeuProThr 460
DB 1429 GGAAGCTTCAGAGGTTCTGGAGATTCCTGAGTGAATCCAGATGCCGCCGCTCCCAAG 1488
QY 461 AspLeuAspaValaGluGlyProTrpPheArgHlsTyAspPheArgGlnSerCystrpVal 480
DB 1489 GACCTGAGAGTGAAGGCGCTTGTTCGSCATTTATGATTTCAACAAGAGCTGCGGCTC 1548
QY 481 ArgAlaIleSerGlnGluaspGlnleuAlaProCystrpGlnAlaGlnHlsProAlaGlu 500
DB 1549 CGTGCCATATCCAGAGAGGACAGCTGGCGCTCTGTGCAAGCTGACACCTCGGGAG 1608
QY 501 ArgValaArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 1609 CCGGAGATGCGGCTTGTGCTCACCGACCTGATTCGACAGAGGACCCCTTTCAGA 1668
QY 521 AlaArgAspaGluGlnProCyAspaAlaProThrSerGlyProCyAlleuCyGlyLeuHlsleu 540
DB 1669 GCTAGGAGCAAGACCGGTGTGCTCCACCTCAGGGGCTTGTGCTGCGGCTCCACCTTG 1728
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1729 GAAAGTTTCAGATTCCTCCAGGCTTC 1755

```

[illegible]

QY 421 ValArgLueApThrTyProValGlyThrgInglyValProSerProAlaLeuAlaGln 440
DB 1949 GTCCGGGAAGACCTACCTACCTGTGGCACTCAGGGGTGTCACGCCGCTGTGCTCAG 2008
QY 441 GlyGlyProGInglySerTPrArgPheLeuGInTPaNSerSerMetProArgLeuProThr 460
DB 2009 GGAGAGACCTCAGGGTTCCTCGAGATTCTCCGACGTGGAATCCATGCCCCGCTCCCAAG 2068
QY 461 AspLeuAspValGluGlyProTPRPhaArgHsTyTAspPheArgGlnSerCySerTPVAl 480
DB 2069 GACCTGGACGTAGAGGGCCCTTGTTCCGCCATTTAGATTTCAGACAGAGCTGTGGGTTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCySerTPrGlnAlaGlnAspProAlaGln 500
DB 2129 COTGCATATCCCAAGAGACCAAGCTGGCCCTCTGTGCAAGCTGGAACACCTCTGGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 2189 CGGGTGAATCGGCTTTCGCTGACCCAGCACTGATTCCGACAGAGGACCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCyValAlaProThrSerGlyProCyLeuCyGlyLeuHsIleu 540
DB 2249 GCTAGGACGACACAGCCGTGTCTCCACCTCAGGGCTTGTGCTGTGCGGCTCCACTTG 2308
QY 541 GluSerSerGlnPheProGlyPhe 549
DB 2309 GAAAGTTCTCACTTCCCTCCAGGCTTC 2335

RESULT 12

ABZ11809
ID ABZ11809 standard; cDNA; 2072 BP.

XX ABZ11809;

DT 20-JAN-2003 (first entry)

DE Human polynucleotide SEQ ID NO 691.

XX Human, genome mapping; gene therapy; food supplement; virus; fungus;
cell-proliferative disorder; neurodegenerative disease; bacterial;
Parkinson's disease; Alzheimer's disease; autoimmune disease;
multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
arthritis; cysticercosis; immunomodulator; neotropic; neuroprotective;
antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
haemostatic; vulnerrary; fungicide; antibacterial; virucide; protozoacide;
antiarthritic; gene; ss.

XX Homo sapiens.

XX OS WO200270539-A2.

XX PD 12-SEP-2002.

XX PF 05-MAR-2002; 2002WO-US005095.

XX PR 05-MAR-2001; 2001US-00799451.

XX (HISE-) HISEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QH, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehman T, Wang J, Wang D, Drmanac RT;
DR WPI; 2002-759812/82.
DR P-PsDB; ABP69592.

XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
XX or coagulation disorders.
XX Claim 1; SEQ ID NO 691; 1012bp + Sequence Listing; English.

CC The invention relates to an isolated polynucleotide (1) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP6902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders, liver
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2072 BP, 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,196-121	Length:	2072
Score:	2827.00	Matches:	525
Percent Similarity:	95.63%	Conservative:	0
Best Local Similarity:	95.63%	Mismatches:	2
Query Match:	94.77%	Indels:	22
DB:	6	Gaps:	1

US-10-071-838-2 (1-549) x ABZ11809 (1-2072)

QY 1 MetAspValAlaGluValAlaGlySerTPRAlaGlnGluArgLueAspIleMet 20
DB 163 ATGACGCTGTAGAGCTGCGCGGCACTTGTTGTGGCCACAAGCGAGACATCATTTATG 222
QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 223 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGACAGAGGGCTTAAGCTTTTCCA 282
QY 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
DB 283 AACTTCAACACACAGCTGCATCATTTGGGATTGTACATAGACGAGCTGCTCTCTG 342
QY 61 ThrAlaArgLueAlaLysGlnIleArgArgLueIleSerArgLysSerTyrTPVAlaAsp 80
DB 343 ACTGGCGGGAGGCGAACAATTCGGCGGAGATCACCGAAGAAGCAAGTGGTGANT 402
QY 81 MetLeuGlyAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 403 ATGCTGGAGAGACTGGAGAAATACAAAGACAGCAAAAGCTCATGATGACGTACAG 462
QY 101 GlyMetProMetAsnIleArgGlyProMetTPrSerValIleLeuAsnIleGluLueMet 120
DB 463 GGAATGCCATGATACATCGGGGCCGAGTGTGTATGTCCTTAAACATAGAGAAATG 522
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerGln 140
DB 523 AAGTTGAAAAACCCCGGAAGATACCATCATGAGAGGAAGGCGACAGGTCATCTGAG 582
QY 141 HisIleGlnArgLysAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 583 CACATCCAGCGCATGACCGGGACCTTAAGCGGACATTAGAGAACATATATCTTCAGG 642
QY 161 AspArgTyrGlyThrLysGlnArgLysLeuLeuHisIleLeuLeuAlaTyrGluLueTyr 180
DB 643 GATCGATACGGAACCAACAGCGGGAACTACTCCATCTCTCTGCTTATAGAGAGTAT 702
QY 181 AsnProGluValGlyTyrCyAspArgAspSerHisIleAlaLeuPheLeuLeuTyr 200
DB 703 AACCGGAGGTGGGTATGCGAGGAGCTGAGCAGCATCGCGCTTGTCTCTCTAT 762
QY 201 LeuProGluGluAspAlaPheTPrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 763 CTTCCTGAGAGAGATCTTCTGCGCACTGGGCACTGGGCAAGTAGAGGACATCC 822

```

QY 221 LeuGlnGlyPheHisSerProAsnGlnGlyThrValGlnGlnGlyLeuGlnAspGlnGlnGln 240
DB 823 CTGACGAGATTTCACAGCCCAATGGCCGAGACGCTCCAGGGGCTCCAAAGCAACAGAG 882
QY 241 HisValValAlaThrSerGlnProIleThrMetGlyHisGlnAspGlySlyAspLeuGly 260
DB 883 CATGTGTAGTACGACGTCACACACCCAAAGACCAATGGGGGATCAG----- 924
QY 261 GlyGlnCysSerProLeuGlyCysLeuLeuIleuLeuLeuAspGlyIleSerLeuGly 280
DB 925 -----ATCTCTCTCGGG 936
QY 281 LeuThrLeuArgLeuLeuTrpAspValTyrIleuValGlnGlnGlnGlnGlnGlnGlnGln 300
DB 937 CTGACCTTGGCCCTGTGGAGCGTGTATCTGTGTAGAGCGCAACAGCGCTTGATCCGATA 996
QY 301 ThrArgIleAlaPheIleValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 320
DB 997 ACAAGAAATCGCTTTAAGTTTCAAGACAGCGCTTCAAGAGACGTCAGGTGTGCGCG 1056
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGlnAspThrValLeu 340
DB 1057 TGGGCACTTTTGGCACTGGCTGTGTGATCTGTGAGCGGCGAGGAGTGAAGACACTGTGCTC 1116
QY 341 LysHisLeuArgAlaSerMetIleSlyLeuThrArgIleGlnGlnGlnGlnGlnGlnGln 360
DB 1117 AACCATCTTGGGCTCTATGAGAGAACTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176
QY 361 AlaIleProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 380
DB 1177 GCCAAACCCGAGACAGAGGCTGTGTGAGCTTCAGGCGCTGTGCGGCTTCACTGTGGGGCT 1236
QY 381 ThrLeuCysLeuGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 1237 ACCCTCTGCAAGGGGAGACAGGAGGCGCTTCAAGGCGCAACAGGCGGCTTCCGCGGCGCC 1296
QY 401 IleTrpSerAlaSerProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
DB 1297 ATTGTGTAGCTTCCCGCCAGCGGAGACCTGTCTTCCACACCTGTCTCTGTGGGGCT 1356
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 1357 GTCCCGGAAAGACACTACCTGTGTGGACACTAGGATGTCCAGGCGCGCTGTGCTCAG 1416
QY 441 GlyGlyProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 460
DB 1417 GGAAGACCTCAGGCTTCTGTGAGATTCCTGACGTGAGACTCCATGCGCGCTCCCAAG 1476
QY 461 AspLeuAspValGlnGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal 480
DB 1477 GACCTGACGATCGAGGGGCGCTTGTGTCCGCAATTATGATTTCAAGACAGAGCTGTGCGCTC 1536
QY 481 ArgAlaIleSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1536
DB 1537 CGTGCATATCCAGAGAGAGACAGCTGCGCTGCTGTGAGGCTGAGCAACCGCTGCGGAG 1596
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlnGlnGlnGlnGln 520
DB 1597 CGGCTGAGATCGGCTTTCGCTCAGCCAGCACTGATTTCCAGCAGGGGACCCCTTCAGA 1656
QY 521 AlaArgAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1656
DB 1657 GCTAAGAGAGAGACAGCGGTGTGTCTCCACTCAGGCGCTTGTCTGTGCGGCTGTGCTTGG 1716
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 1717 GAAAGTCTCAGTTCCTCCAGGCTTC 1743

```

RESULT 13

ADM44327

ADM44327 standard; cDNA, 2072 BP.

AC ADM44327;

```

XX 03-JUN-2004 (first entry)
DT
XX
XX
DE Novel human arginine-rich protein cDNA #691.
KW ss; gene; human; arginine-rich protein; cancer; inflammation;
XX genetic disorder.
XX Homo sapiens.
XX US2004053250-A1.
XX 18-MAR-2004.
XX 21-NOV-2002; 2002US-00302172.
XX 05-MAR-2001; 2001US-00799451.
XX 05-MAR-2002; 2002WO-US005095.
XX 20-AUG-2002; 2002US-00225251.
XX (TANG/) TANG Y T.
XX (XUEA/) XUE A.
XX (DRMA/) DRMANAC R T.
PI Tang YT, Xue A, Drmanac RT;
DR WPI; 2004-238579/22.
XX
XX PT New isolated arginine-rich protein-like polynucleotides and polypeptides,
XX PT useful for diagnosing and/or treating conditions associated with aberrant
XX PT activity of the arginine-rich polypeptides, such as cancer and
XX PT inflammation.
XX PS Disclosure, SEQ ID NO 691; 51bp; English.
XX
XX CC The invention relates to an isolated polynucleotide. The methods and
XX CC compositions of the present invention are useful for the diagnosis and/or
XX CC treatment of diseases or conditions associated with aberrant expression
XX CC or activity of the arginine-rich protein-like polypeptides, such as
XX CC cancer and inflammation. They can also be used in forensic, gene
XX CC mapping, identification of mutations responsible for genetic disorders,
XX CC and in assessing biodiversity. The present sequence represents a novel
XX CC human arginine-rich protein cDNA.
SQ Sequence 2072 BP; 525 A; 568 C; 577 G; 402 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,19e-121
Score: 2827.00 Length: 2072
Percent Similarity: 95.63% Matches: 525
Best Local Similarity: 95.63% Conservative: 0
Query Match: 94.77% Mismatches: 2
DB: 12 Indels: 22
Gaps: 1
US-10-071-838-2 (1-549) x ADM44327 (1-2072)
QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGlnGlnGlnGlnGlnGlnGln 20
DB 163 ATGACCTGTGTAGAGGTGCGCGGACAGTTGGTGGCAACAGCGAGAGACATCAATTATG 222
QY 21 LysTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 222
DB 223 AAATACGAAAAGGAGACACGAGCTGGGCTGCCAGAGGACAGGCGCTTAACCTTTTGA 282
QY 41 SerTrpAsnAsnAsnValAspHisLeuGlyIleValHisGlnThrGlnLeuProLeu 60
DB 283 AGCTACAAACAAACATCATCATTTGGGATTTGACATGACAGGACGCTCTCTTG 342
QY 61 ThrAlaArgGlnAlaIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
DB 343 ACTGCGCGGAGGCGCAAGCAATTCGCGGAGATCAGCCGAAAGACCAAGTGGTGGAT 402
QY 81 MetLeuGlyAspTrpGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 100

```

```

Db      403 ATGCTGGGAGATGGGAGAAATACAAAGCAGCAAGAAAGCTCATGATGACGATACAG 462
Qy      101 GYMETPROMETAEMLIETARGIYPROCHETTPSERVALLEULEUBENILEGLUWET 120
Db      463 GGAATGCCCATGATCCGGGGCCGATGTGTGTCATGCTCTTCAACCTGAGAAATG 522
Qy      121 LYSLEULVAAANPROGLIYARGTGTINILEMETLVSGIULVYGAIRGSESGI 140
Db      523 AAGTTGAAACCCCGAGAGATCCAGATCATGAGAGAGAGAGAGAGAGAGATCTGAG 582
Qy      141 H4SILEGINARGLIETAPARSHAPVALSERGIYTHLEUARGIYH4SILEPHEARG 160
Db      583 CACATCCAGCGCATGACCGGGACGTAAGCGGACATTAAGAACATATATCTTCAGG 642
Qy      161 AEPARGIYGLIYTHLYEGINARGLIULEULEUH4SILEULEUALATYGLIULYR 180
Db      643 GATCATATCGGAAACAGACGCGGAACTACTCCATCTCTCTGCAATATGAGAGAT 702
Qy      181 AANPROGIUVALIYTYCYEARSHAPLEUSERH4SILEALALEUPHEULEUTYR 200
Db      703 AACCGAGAGTGGCTACTGACAGGACCTGAGCCACATGCGCTGTGCTCTCTAT 762
Qy      201 LEUPROGIUGIUAAPALAPHEITPALEUVALIGIULEULEUALSERGIUARGHISER 220
Db      763 CTTCCTGAGAGAGATGATCTGCGGCACTGTCGACGCTGCGCAGTGAAGGACATCC 822
Qy      221 LEUGINGIYPHEHISERPROBANGIYGLIYTHVALIGINGIYLEUGINAPGINGINU 240
Db      823 CTGCGAGGATTTCAACAGCCCAATGCGGAGCCGTCACAGGGGCTCCAGAACCAAGAG 882
Qy      241 H4SVALIVALIATHSERGIUPROLYERTHMETGIYH4SINAPLYBLYSHAPLEUCYS 260
Db      883 CATGGGATGACCATCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 924
Qy      261 GLIYGINYSERPROLEUGIYCYLEULIETARGILEULEIETASPGIYIIESERLEUGI 280
Db      925 -----ATCTCTCGGG 936
Qy      281 LEUTHLEUARGLEUTRPAAPVALIYRIEUVAILIGIUGIYGLIUALALEUWETPROILE 300
Db      937 CTCACCTCGCGCTGTGGGACCTGTATCTGTAGAAAGGAGCAAGCGCTGATGCGAT 996
Qy      301 THRARGILEALAPHELYVALIGINGIYSAARGLEUTHLYTHSERARGYVGLIYPRO 320
Db      997 ACAAAATCGCTTTAAGGTTCAGCAAGAGCGCTTCACAAACGTCAGGAGTGGCCG 1056
Qy      321 TRPALAARGPHECYASNAARGPHEVALIAPTHRTPALAARGAPGLIUAAPTHRVALLEU 340
Db      1057 TGGGACCTTTTGGCAACCGCTTCTGTATCTGCGGCAAGGAGAGACCTGTGCTC 1116
Qy      341 LYSHISLEUARGALASERMETLYALYSLEUTHARGLYVGLIYASPLEUPROBPRO 360
Db      1117 AAGCATCTTAGGGCTCTATAGAAACTTAACAAAGAGAGAGAGAGAGAGAGAGAGAG 1176
Qy      361 ALALYSRPROGIUGINGIYSSERVALSERARGPROVALIAPSERARGIYGLIYYS 380
Db      1177 GCCAAACCCGAGCAAGGCTGTCGATCCAGGCTGTGCGCTTCACGTCGCGAGAG 1236
Qy      381 THRLEUCYLYSGIYASPARGLIAPPROFGLIYPROFGLIYARGPHEPROARGPRO 400
Db      1237 ACCCTTCGAAAGGGGAGACAGGAGCCCTCCAGGCCACAGCCCGGTCCCGGCGCC 1296
Qy      401 ILETSPERALASERPROPROARGALAPROARGSERSETHPROCYSPROGLIYGLIYALA 420
Db      1297 ATTTGTGAGCTTCCCGCACGCGGACCTGTTCTTCACACCTGTGCTGTGTGGGCT 1356
Qy      421 VALARGIYUAPRTHRTYRPROVALIYTHRINGIYVALIAPROSERPROALALEUVALGIN 440
Db      1357 GTCCGGAGAGACCTACCTGTGTGGCACTCAGGGGTGTGCCAGCCGCGCTGCTCAG 1416
Qy      441 GLIYGIYPROGINGIYSEITPARPHELEUGINTPASENSEMEPROARGLEUPROTHR 460

```

```

Db      1417 GGAGACCTCAGGGTTCCTGAGATTCCTGCAGTGAACTCCATGCCCGCCTCCAAAG 1476
Qy      461 ASPLEUAPVALIGIUGIYPROTPRPHARGH4SITPAAPPEARXGINSERCYSTRPAL 480
Db      1477 GACCTGGAGCTAGAGGGCCCTTGCTCCGCAATATGATTCAGACAGAGCTGCTGGCTC 1536
Qy      481 ARGALAIIESERGIUGIUAAPGLIULEUALAPROCYSTRPGLIUALIAGIYH4SIPROALIGU 500
Db      1537 CGTGCATATCCAGAGAGACCAAGCTGCCCCCTGCTGAGAGCTGAAACACCTCGGAG 1596
Qy      501 ARGVALIATSERVALAPHEALAPALAPROSERTHRAPSERAPGINGIYTHPROPHARG 520
Db      1597 CGGGTGAATCGGCTTTGTGCTGACCCAGCACTGATTCGACAGGGGACCCCTTCAGA 1656
Qy      521 ALAARGAPGIUGIINPROCYALAPROTHSERGIYPROCYBLEUCYAGIYLEUHISLEU 540
Db      1657 GCTAGAGAGCAACAGCCGTGTGCTCCACTCAGAGGACCTGTGCTGTGCGCTCCACTTG 1716
Qy      541 GUSERSERGIUPHEPROPROGLIYPHE 549
Db      1717 GAAGTTCTCAGTTCCCTCCAGGCTTC 1743

RESULT 14
AD37386
ID AD37386 standard; DNA, 2647 BP.
AC AD37386;
XX
XX
XX 18-DEC-2003 (first entry)
DE
XX
XX Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 219.
KW
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KW immunomodulator; Cyclostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiac; Gene therapy; human; gene; de.
XX
XX Homo sapiens.
OS
XX
XX WO2003048202-A2.
XX
XX 12-JUN-2003.
XX
XX 03-DEC-2002; 2002MO-JP012644.
XX
XX 03-DEC-2001; 2001JP-00368692.
XX
XX 05-DEC-2001; 2001US-0335829P.
XX
XX 03-OCT-2002; 2002JP-00291302.
XX
XX 04-OCT-2002; 2002US-0415769P.
XX
XX (ASAH) ASAH KASEI KK.
XX
XX Matsuda A, Muramatsu S;
XX
XX MPI: 2003-505282/47.
XX
XX P-PSDB; AD37387.
XX
XX
XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
XX useful for treating inflammation, autoimmune diseases, cancers,
XX infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX ischemic disorders.
XX
XX Claim 4; SEQ ID NO 219; 938bp; English.
XX
XX
XX The present invention relates to novel proteins and their coding
XX sequences (AD37386-AD37455), which activate nuclear factor kappa B (NF-
XX kappaB). The proteins and their coding sequences are useful for treating
XX a disease associated with NF-kappaB activation, such as inflammation,
XX autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
XX neurodegenerative diseases, or ischemic disorders.
XX
XX
XX Sequence 2647 BP; 609 A; 765 C; 750 G; 523 T; 0 U; 0 Other;
XX
XX

```

Alignment Scores:

Pred. No.: 2.76e-121 Length: 2647
 Score: 2827.00 Matches: 525
 Percent Similarity: 95.63% Conservative: 0
 Best Local Similarity: 95.63% Mismatches: 22
 Query Match: 94.77% Indels: 2
 Gaps: 1

US-10-071-838-2 (1-549) x ADC37386 (1-2647)

```

QY 1 MetAspValValGluValAlaGlySerTrpTrpAlaGlnGluValArgGluAspGluLeuMet 20
Db 755 ATGAGCGTGTGATAGAGGTCCGGCGCAGTGTGTGGGACAGAGCCAGAGGACATCATTAATG 814
QY 21 LysTrpGluLysGlyHisArgAlaGlyLeuProGluLysAspGlyProLysProPheArg 40
Db 815 AAATACGAAAGGACACCGAGCTGGCTGCCAGAGGACAGAGGGCTTAAGCTTTTCA 874
QY 41 SerTrpAspAspAspValAspHisLeuGlyTlleValHisGluThrGluLeuProPheLeu 60
Db 875 ACCTCAACACACACAGTGCATCATTTGGGATTTGATGAGACGAGCTGCTCTCTG 934
QY 61 ThrAlaArgGluAlaLysGlnLeuArgGluLysSerArgLysSerLysTrpValAsp 80
Db 935 ACTGGCGGAGGAGGAGCAAAATTCGGCGGAGATCAGCCGAAAGGACAGTGGGTGAT 994
QY 81 MetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuLysAspArgAlaLys 100
Db 995 ATGCTGGGAGACCTGGAGAAATACAAAGACAGACAGAGCTCATGATCGAGCTTACAG 1054
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluMet 120
Db 1055 GGAATGCCATGAAATCCGGGCGCCGATGTGTCATCTCTGAACTGAGGAATG 1114
QY 121 LysLeuLysAspProGlyArgTrpGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
Db 1115 AAGTTGAAAAACCCGGAGAGATACCAAGATCAGAGAGGACAGAGGTGATCTGAG 1174
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 1175 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTAGAGAACATATATCTTCAG 1234
QY 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluLys 180
Db 1235 GATCGATACGGAAACCAAGCAGCGGAACTCTCACATCTCTGCGATATAGAGAT 1294
QY 181 AsnProGluValGlyTrpArgAspLeuSerHisIleAlaLeuPheLeuLeuTrp 200
Db 1295 AACCGAGGTGGGCTACTGACAGGACCTGAGCCATCGCGCTTGTCTCTCTAT 1354
QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGluLeuLeuAlaSerGluArgHisSer 220
Db 1355 CTTCCTGAGAGAGATGATCTGGGCACTGGTGCAGCTGTGGCCAGAGAGGACCTCC 1414
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 1415 CTCAGAGGANTTCAACGCCCAATATGCGGAGCCGCTCCAGGGCTCAAGACCAACAGAG 1474
QY 241 HisValValAlaThrSerGlnProLysTrpMetGlyHisGlnAspLysLysAspLeu 260
Db 1475 CATGTGTAGCCACGTCACCAACCCAGACCATGGGGCATCG----- 1516
QY 261 GlyGlnLysSerProLeuGlyCysLeuLysArgIleLeuLysAspGlyLysSerLeu 280
Db 1517 -----ATCTCTCTCGGG 1528
QY 281 LeuThrLeuArgLeuTrpAspValTrpLeuValGluGlyGluAlaLeuMetProIle 300
Db 1529 CTCACCTTGCGCTGTGTGAGCGTGTATCTGTGAGAGCGAAACAGCGCTGATGCGGATA 1588
QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320

```

```

Db 1589 ACAGAAATCGCTTTAAGTTTCAGCAGAAAGCGGCTTCACGAAAGCTCCAGAGTGGCCCG 1648
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
Db 1649 TGGCAGCGTTTTCGAAACCGGTTCTGTGATACCTGGCCAGGAGATGAGACATGTGTCTC 1708
QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1709 AAGCATTTTGGGCTCTTATGAAAGAACTAACAGAAAGAGGGGACCTTCACACCCCA 1768
QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1769 GCCAAACCCGAGCAGAGGTCGTCCGACATCCAGGCTGTGGCGCTTCACGTGGCGGAG 1828
QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGAAAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1888
QY 401 IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysAspGlyGlyAla 420
Db 1889 ATTGTGTACGCTTCCCGCAGCGGACCTGCTTTTCAACACCTGTCTGTGTGGGCT 1948
QY 421 ValArgGluAspTrpTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCCGGAGAGACACTTACCTGTGTGGGAGCTCAGGGGTGTGCCAGCCGGCCCTGGCTCAG 2008
QY 441 GlyGlyProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr 460
Db 2009 GGAAGACCTCAGAGGTTCTCTGAGATTCCTGAGTGAATCTCAAGTCCCGCTCCACAG 2068
QY 461 AspLeuAspValGluGlyProTrpPheArgHisTrpAspPheArgGlnSerCysTrpVal 480
Db 2069 GACCTGAGCTAGAGGGCTTGTGTTCCGCAATTAATGATTTCAACAGAGACTGTGGGTC 2128
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAlaGlu 500
Db 2129 CGTGCAATATCCAGAGAGGACAGCTGCGCCCTGCTGGAGCTGAACACCTTGGCGAG 2188
QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
Db 2189 CGGTGTAGATCGGCTTGTGCTGCACCAAGCATGATTCGACACAGGGCACCCCTTCAGA 2248
QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuGlyGlyLeuHisLeu 540
Db 2249 GCTAGAGACAAACGCGTGTGCTCCACTCAGGGCTTGTGCTGTGGGCTTCCACTTG 2308
QY 541 GluSerSerGlnPheProProGlyPhe 549
Db 2309 GAAAGTTCTCAGTTCCCTCCAGGCTTC 2335

RESULT 15
ADM01991
ID ADM01991 standard; cDNA, 2647 BP.
XX
AC ADM01991;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human cDNA of the invention SEQ ID NO:676.
XX
KM ser, gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
OS Homo sapiens.
XX
FN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
PA (REAS-) RES ASSOC BIOTECHNOLOGY.

```

XX Isegai T, Sugiyama T, Otsuki T, Makamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 DR WPI: 2003-723558/69.
 DR P-PSDB: ADM04434.

PT New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 676; 305bp; English.

XX The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC cDNA sequence of the invention.

XX Sequence 2647 BP; 609 A; 765 C; 750 G; 523 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,766-121 Length: 2647
 Score: 2827.00 Matches: 525
 Percent Similarity: 95.63% Conservative: 0
 Best Local Similarity: 95.63% Mismatches: 2
 Query Match: 94.77% Indels: 22
 DB: 11 Gaps: 1

US-10-071-838-2 (1-549) x ADM01991 (1-2647)

QY 1 MetAaPValIValIGluValAlaGlySerTPRTPrAlaGlnGluArgGluAspIleIleMet 20
 DB 755 ATGAGCGTGTGAGAGCTCCGGGCACTTGTGTGGGCAAGACGAGGACATCTATTATG 814
 QY 21 LysTyrGluLysGlyVHIsArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 DB 815 AAATACGAAGAAGGACACCGAGCTGGGCTGCCAGAGCAAGAGGGGCTTAAGCTTTTGA 874
 QY 41 SerTyrAaAsnAsnValAspHIsLeuGlyIleValHIsGluThrGluLeuProPheLeu 60
 DB 875 AGCTACAAACAACAGCTGATCTTGGGGATTGTATCATGAGACGAGCTCTCTCTG 934
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysSerValAsp 80
 DB 935 ACTGGCGGGAGGCGAAGAAATTCGGCGGGAGATCAGCCGAAGAAGCAAGTGGGTGGAT 994
 QY 81 MetLeuGlyAspTPrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
 DB 995 ATGCTGGGAGACTGGGAGAAATACAAAGCAAGAGCTCATGATGAGCGTATCAAG 1054
 QY 101 GlyMetProMetAsnIleArgGlyProMetTPrSerValLeuLeuAsnIleGluGluMet 120
 DB 1055 GGAATGCCCATGAAATCCGGGGCCCGATGTGTCTCACTCTCTAACAACCTGAGGAAATG 1114
 QY 121 LysLeuLysAsnProGlyVArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 DB 1115 AAGTTGAANAACCCCGAAGATACCAATCATGAAGAGAGAGGCAAGAGGCTCATCTGAG 1174
 QY 141 HisIleGlnArgGlyLeuAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 DB 1175 CACATCCAGCGCATGACCGGAGCGTAAAGCGGACATTAAAGAAACATATATCTTCAAG 1234
 QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGluTyr 180
 DB 1235 GATCATATCGGAACCAAGACGCGGAACTACTCAATCTCTCTGCGATATAGAGATAT 1294

QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleLeuLeuLeuTyr 200
 DB 1295 AACCCGAGAGTGGGCTATCGAGGAGCACTGAGCCACATCGCGCTTGTCTCTCTAT 1354
 QY 201 LeuProGluGluAspAlaPheTPrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 DB 1355 CTTCCTGAGGAGATGATCTCGGCACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1414
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
 DB 1415 CTGACGAGATTCACAGCCCAATGCGGAGCCGCTCCAGGGGCTCCAGACCAAGAGAG 1474
 QY 241 HisValAlaAlaThrSerGlnProLysThrMetGlyHisIleAspLysLysAspLeuCys 260
 DB 1475 CATGTGTAGGACAGCTCAACCCCAAGCATGGGAGCTCAG-----ATCTCTCTCGGG 1516
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyTyrIleSerLeuGly 280
 DB 1517 -----ATCTCTCTCGGG 1528
 QY 281 LeuThrLeuArgLeuTPrAspValTyrLeuValGlnGlyGlnGlnAlaLeuMetProIle 300
 DB 1529 CTCACCTCGCGCTGTGGAGAGTGTATCTGTGAGAGGCGAAGCAGCGCTTATATGCCATTA 1588
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 DB 1589 ACAAGATCGCTTTAAAGTTTACAGAGAGGCGCTCAAGAGAGCTCCAGAGGTGTGGCCG 1648
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTPrAlaArgAspGluAspThrValLeu 340
 DB 1649 TGGGCACTTTTGGCAACCGGCTTCTTATATCTGGGCGAGGAGTACAGACCTGTGCTC 1708
 QY 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
 DB 1709 AAGCATCTTAGGCTCTATAGAACTAACAAGAGAGAGGAGGAGCTTGCACACCCCA 1768
 QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLysLys 380
 DB 1769 GCCAAGCCGAGCAAGAGGCTGTGCGCATCCAGGCGCTTCAAGCGGGGGAAG 1828
 QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 DB 1829 ACCCTCTGCAAGGGGAGAGAGGAGGCGCTCAGGCGCCACAGAGCGCGGCTCCGCGGCC 1888
 QY 401 IleTPrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyLysAla 420
 DB 1889 ATTGTGATGCTTCCCGCCAGGCGACCTGTTCTTCAACCTGTGCTGTGGTGGCT 1948
 QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 DB 1949 GTCCGGAGAGACACTTACCTGTGGCACTCAAGGTGTGCCAGCGCGGCTGCTCAG 2008
 QY 441 GlyGlyProGlnGlySerTPrArgPheLeuGlnTPrAsnSerMetProArgLeuProThr 460
 DB 2009 GAGAGACCTCAGGGGTCTCGAGATTCCTGAGATGAACTCCATCCCGCTCCCAAG 2068
 QY 461 AspLeuAspValGluGlyProTPrPheArgHisTyrAspPheArgGlnSerCysTyrVal 480
 DB 2069 GACCTGAGCGTGAAGAGGCGCTTGTGCGCAATTATGATTCAGCAAGAGCTGCTGGGCT 2128
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTPrGlnAlaGluHisProAlaGlu 500
 DB 2129 COTGCAATTCCTCAGAGAGACCAAGTGGCCCTGCTGCTGAGCGTGAACACCTGGGAG 2188
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 DB 2189 CGGGTGAATCGGCTTGTGCTGCAACCGACCTGATTCGACCAAGGAGCAACCCCTTCAAG 2248
 QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisIleu 540
 DB 2249 GCTAGGAGCAAGAGCGCTGTGCTCCACTCAGAGGCTTGTGCTGCGGCTTCACTTG 2308
 QY 541 GluSerSerGlnPheProGlyPhe 549

Mon Feb 7 06:58:10 2005

us-10-071-838-2.rng

Page 20

Db 2309 GAAGTTCTCAGTTCCTCCAGGCTTC 2335

Search completed: February 4, 2005, 08:20:27
Job time : 819 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 10:14:14 ; Search time 853 Seconds

(without alignments)
3698.116 Million cell updates/sec

Title: US-10-071-838-2
Perfect score: 2983
Sequence: 1 MDVAVGSMWAQEREDIM.....TSGPCLGLHSSQPPPGF 549

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.epool/US10071838/runat_03022005_071221_304/app_query.fasta_1.711
-DB=Published Applications NA -OFMT=fastcap -SUFFIX=rnph -MINMATCH=0.1
-LOOCL=0 -LOOEXT=0 -UNITS=bite -START=1 -END=1 -MATRIX=10sum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=spc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10071838_@CNS_1_1_627_@runat_03022005_071221_304
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	1964	14 US-10-071-838-1	Sequence 1, Appl
2	2978	99.8	1993	18 US-10-839-882-31	Sequence 31, Appl
3	2907.5	97.5	2146	14 US-10-071-838-3	Sequence 3, Appl
4	2850	95.5	2084	16 US-10-302-172-692	Sequence 692, App
5	2827	94.8	2072	16 US-10-302-172-691	Sequence 691, App
6	2827	94.8	2647	16 US-10-108-260A-676	Sequence 676, App
7	2784	93.3	1752	15 US-10-094-466-61	Sequence 61, Appl
8	2773	93.0	1862	14 US-10-071-838-5	Sequence 5, Appl
9	2191.5	73.5	7878	9 US-09-962-436-562	Sequence 562, App
10	2133.5	71.5	8201	17 US-10-283-975A-272	Sequence 272, App
11	2133.5	71.5	8408	13 US-10-098-841-51	Sequence 51, Appl
12	2133.5	71.5	8420	13 US-10-098-841-50	Sequence 50, Appl
13	1507.5	50.5	9805	9 US-09-764-869-1824	Sequence 1824, App
14	1507.5	50.5	9805	10 US-09-764-891-1857	Sequence 7857, App
15	1507.5	50.5	9805	14 US-09-091-504-1824	Sequence 1824, App
16	1507.5	50.5	9805	16 US-10-227-577-1824	Sequence 1824, App
17	1500.5	50.3	9404	9 US-09-764-869-1826	Sequence 1826, App
18	1500.5	50.3	9404	10 US-09-764-891-1825	Sequence 7859, App
19	1500.5	50.3	9404	14 US-10-091-504-1825	Sequence 1826, App
20	1500.5	50.3	9404	16 US-10-227-577-1826	Sequence 1826, App
21	1469	49.2	9804	9 US-09-764-869-1829	Sequence 1829, App
22	1469	49.2	9804	10 US-09-764-891-1861	Sequence 7861, App
23	1469	49.2	9804	14 US-10-091-504-1829	Sequence 1829, App
24	1469	49.2	9804	16 US-10-227-577-1829	Sequence 1829, App
25	1465	49.1	9796	9 US-09-764-869-1825	Sequence 1825, App
26	1465	49.1	9796	10 US-09-764-891-1858	Sequence 7858, App
27	1465	49.1	9796	14 US-10-091-504-1825	Sequence 1825, App
28	1465	49.1	9796	16 US-10-227-577-1825	Sequence 1825, App
29	1428.5	47.9	6696	9 US-09-764-869-1821	Sequence 1821, App
30	1428.5	47.9	6696	10 US-09-764-891-1856	Sequence 7856, App
31	1428.5	47.9	6696	14 US-10-091-504-1821	Sequence 1821, App
32	1428.5	47.9	6696	16 US-10-227-577-1821	Sequence 1821, App
33	1355.5	45.4	9831	9 US-09-764-869-1827	Sequence 1827, App
34	1355.5	45.4	9831	14 US-10-091-504-1827	Sequence 1827, App
35	1355.5	45.4	9831	16 US-10-227-577-1827	Sequence 1827, App
36	1355	45.4	3979	9 US-09-764-869-1828	Sequence 1828, App
37	1355	45.4	3979	14 US-10-091-504-1828	Sequence 1828, App
38	1355	45.4	3979	16 US-10-227-577-1828	Sequence 1828, App
39	1053	35.3	886	9 US-09-764-869-1831	Sequence 1831, App
40	1053	35.3	886	14 US-10-091-504-1831	Sequence 1831, App
41	1053	35.3	886	16 US-10-227-577-1831	Sequence 1831, App
42	1053	35.3	886	14 US-10-091-504-1832	Sequence 1832, App
43	1053	35.3	886	16 US-10-227-577-1832	Sequence 1832, App
44	1053	35.3	886	16 US-10-227-577-1832	Sequence 1832, App
45	818	27.4	470	9 US-09-815-343-501	Sequence 501, App

ALIGNMENTS

RESULT 1
US-10-071-838-1
Sequence 1, Application US/10071838
Publication No. US20030044814A1
GENERAL INFORMATION:
APPLICANT: Li, Jing
APPLICANT: Powers, Scott
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: PCT/7. An Amplified Cancer Gene
FILE REFERENCE: 018781-007610US
CURRENT APPLICATION NUMBER: US/10/071.838
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: US 60/267,615
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1964
TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: human PRC17
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(1650)
 US-10-071-838-1

Alignment Scores:

Pred. No.:	1,086-265	Length:	1964
Score:	2983.00	Matches:	549
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-071-838-2 (1-549) x US-10-071-838-1 (1-1964)

```

QY      1 MetAPValValAGValAlAGlySerTTPAlAGInGUaArgGUaAPVleIleMet 20
DB      1 ATGACCTGTAGAGGTGCGCGGCACTGGTGGGCAAGAGCGAGAGGACATCATATG 60
QY      21 LysTyrGluYsgIyHlSARGalAGlyLeuProGUaAPlyGlyProlyAProPheArg 40
DB      61 AAATACAAAAGGAGCACCGAGCTGGCTGCCAGAGGACAAAGGGGCTTAAGCTTTTCA 120
QY      41 SerTyrAspAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProleu 60
DB      121 AGCTACACAAACAACCTTCATCTTTGGGATTTGATCATGAGACGGAGCTGCTCTCTG 180
QY      61 ThrAlaArgGUaAlaYsgInUleArgArgGluLeuSerArgYsSerlySTTPValAsp 80
DB      181 ACTGCGCGAGAGCGAGGACAAATTCGGCGGAGATCCGCGAGACCGAAGACAAAGTGGTGA 240
QY      81 MetLeuGlyAspTTPGluYleTyrYsSerSerArgYsLeuIleAspArgAlaTyrIys 100
DB      241 ATGCTGGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCATAGTCAGGCTACAG 300
QY      101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
DB      301 GGAATGCCCATGAAATATCCGGGCGCCGATGGTGCATGCTCTCTGAACTAAGGAAATG 360
QY      121 LysLeuYsAsnProGUaArgYsArgYsInUleMetYsGlyYsArgSerSerGlu 140
DB      361 AAGTTGAAAACCCCGGAGATACAGATCATAGAGAGAGAGGAGGACAGGCTCATCTGAG 420
QY      141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgYsHisIlePhePheArg 160
DB      421 CACATCCAGGCAATCAACCGGAGCGGACGTACCGGAGCATTAAGAGCANTATTTCTCAGG 480
QY      161 AspArgTyrGlyThrIysGlnArgGluLeuLeuHisIleLeuAlaIlyrGluGluTyr 180
DB      481 GATCGATACGAAACCAAGACGCGGACATCTCCACATCTCTCGGCAATGAGAGAT 540
QY      181 AsnProGUaValGlyTyrCysArgAspLeuSerHisIleAlaIleuPheLeuLeuTyr 200
DB      541 AACCCGAGAGGTGGCTACTCGAGGAGACTGAGCCACATCCGCCCTTGTTCTCCCTCAT 600
QY      201 LeuProGUaAspAlaPheTTPAlaLeuValGluLeuLeuAlaSerGluThrHisSer 220
DB      601 CTTCCTGAGAGAGATGATCTCTGGGCACTGGTGAAGTGTGCTGAGTAAAGGCACTCC 660
QY      221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
DB      661 CTGACGGGATTTTCACAGCCCAATGCGGAGCGGCTCAGAGGGCTCCCAAGACCAAGAG 720
QY      241 HisValValAlaThrSerGlnProlySTTPMetGlyHisGlnAspYsLeuAspLeu 260
DB      721 CATGTGTACCAAGCTACCAACCAAGACCATGGGCAATCGAGACAAAGAAAGATCATGT 780
QY      261 GlyGlnCysSerProleuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
DB      781 GGGCAGATGTTCCCGTTAGGCTGCTCATCCGGAATATGATGACGGGATCTCTCGGG 840
  
```

```

QY      281 LeuThrLeuArgLeuTTPAspValTyrIleuValGlnGlyGlnAlaIleuMetProIle 300
DB      841 CTCAACCTGGCGCTGTGGAGCGTATCTGTAGAGAGGAGGACAGGCTTATGATCCCAT 900
QY      301 ThrArgIleAlaPheYsValGlnGlnIySARGalAGlyLeuThrYsThrSerArgYsGlyPro 320
DB      901 ACAGAGATGCCCTTTAAGGTTCAGCAGAAAGCGCTCAAGAAAGCTTCAGGTGGCCG 960
QY      321 TTPAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspGluAspThrValLeu 340
DB      961 TGGCAGCGTTTTCACACCGGTTCTTATACCTGGGCGAGGAGTGAACACTGTGCTC 1020
QY      341 YsHisIleuArgAlaSerMetIlySleuThrArgYsGlnGlyAspLeuProProPro 360
DB      1021 AAGCATCTTAGGCGCTCTATAGAAACCTAACAAAGAAAGAGGAGACCTGCAACCCCA 1080
QY      361 AlaYsProGUaGlnGlySerSerAlaSerArgProValProAlaSerArgIyGlyIys 380
DB      1081 GCCAAACCCGAGCAAGGCTGTGGCATCCAGGCTGTGCGGCTTCACTGGCGGGAAG 1140
QY      381 ThrLeuCysYsGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB      1141 ACCCTCTGCAAGGGGAGACAGGAGGCTGGCATCCAGGCTGTGCGGCTTCACTGGCGGGAAG 1140
QY      401 IleTSPSerAlaSerProProArgAlaProArgSerSerThrProCysPProGlyValAla 420
DB      1201 ATTGTGACACTTCCCGGACAGGAGCACTGCTTTCACACCTGTCTGTGGGGGCT 1260
QY      421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB      1261 GTCGGGAGAGACACCTACCTGTGGGCACTAGGAGTGTGCCACCGGCGGCTGGGCT 1320
QY      441 GlyIyProGlnGlySerTTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
DB      1321 GAGAGACCTCAGAGGTTCTGTAGATTCCTGACGTGAGAACCTATGCCGCCGCCCAAG 1380
QY      461 AspLeuAspValGlnGlyProTTPPheArgHisTyrAspPheArgGlnSerCysTTPVal 480
DB      1381 GACCTGAGCGTAGAGGGGCTGTGGTCCGCACTTATGATTTACAGACAGCTGCGGCT 1440
QY      481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTTPGlnAlaGlnIleProAlaGlu 500
DB      1441 CGTGCATATCCAGAGAGAGACAGCTGGGCCCTGCTGCAAGGCTGACACCTCGGAG 1500
QY      501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB      1501 CGGAGTGAATCGGCTTTCCTGCACTGCACTGATTCGACAGGAGGACCCCTTCAGA 1560
QY      521 AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisIleu 540
DB      1561 GCTAGGAGAGACAGCGGCTGTCTCCACCTCAGAGGCTGCTGCTGCGGCTTCACCTG 1620
QY      541 GluSerSerGlnPheProProGlyPhe 549
DB      1621 GAAAGTTCACATCTCTCCACAGGCTTC 1647
  
```

RESULT 2

US-10-839-882-31
 Sequence 31, Application US/10839882
 Publication No. US20040203106A1
 GENERAL INFORMATION:
 APPLICANT: INCYTE PHARMACEUTICALS, INC.
 APPLICANT: TANG, Y.-Tom
 APPLICANT: YUE, Henry
 APPLICANT: HILLMAN, Jennifer L.
 APPLICANT: GUEGLER, Karl J.
 APPLICANT: CORLEY, Neil C.
 APPLICANT: LAI, Preeti
 APPLICANT: AZIMZAI, Yaida
 APPLICANT: BAUGHN, Mariah R.
 APPLICANT: JUNMING, Yang
 APPLICANT: SHIH, Leo L.

TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
FILE REFERENCE: PF-0619 PCT
CURRENT APPLICATION NUMBER: US/10/839,882
CURRENT FILING DATE: 2004-05-05
PRIOR APPLICATION NUMBER: US/09/807,452
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 09/175,737, unassigned; 60/118,559; 09/249,740; unassigned;
60/154,336
PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11;
1999-04-22
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Perl Program
SEQ ID NO: 31
LENGTH: 1993
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
FEATURE:
OTHER INFORMATION: Incyte ID No: 4325626CB1
US-10-839-882-31

Alignment Scores:

Pred. No.:	3,196-265	Length:	1993
Score:	2978.00	Matches:	548
Percent Similarity:	99.82%	Conservative:	0
Best Local Similarity:	99.82%	Mismatches:	1
Query Match:	18	Indels:	0
DB:		Gaps:	0

US-10-071-838-2 (1-549) x US-10-839-882-31 (1-1993)

1 MetAaPvAlValGluValAlaGlySerTPPTAlaGlnGluArgLysAlaPheMet 20
42 ATGAGAGTGTAGAGTCCGGGACAGTGTGGGCAAGAGGAGGAGCATCATTAAG 101
21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
102 AATAGCAAGAGGACACCGAGCTGGGCTGCCAGAGACAGGGGCTTAAGCTTTTCA 161
41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
162 ACCTCAACAACAACCTCATCATTTGGGGATTGTTCATGACAGCAGCTGCTCTCTG 221
61 ThrAlaArgGluAlaLysGlnIleArgGluIleSerArgLysSerLysTyrValAsp 80
222 ACTGCGCGGAGCGGAGCAAAATTCGCGGAGATCAGCCGAAAGCAAGTGGTGAT 281
81 MetLeuGlyAspTyrGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
282 ATGCTGGAGACTGGGAGAAATACAAAGACAGCAAAAGCTCATGATGAGGCTACAG 341
101 GlyMetProMetCAsnIleArgGlyProMetTyrSerValLeuLeuAsnIleGluGlnMet 120
342 GGAATGCCCATGAACATCCGGGGCCGAGATGGTGCATCTCTGAAACACTGAGGAATG 401
121 LysLeuLysAsnProGlyLysTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
402 AAGTTGAAACCCCGGAGATACAGATCATGAAGAGAGAGGCAAGGCTATCTGAG 461
141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
462 CACATCCAGCATGCAACCGGAGCTTAACGGGACATTAAGAGCATATATCTTCAAG 521
161 AspArgTyrGlyThrLysGlnArgLysLeuLeuHisIleLeuLeuAlaTyrGluGlnTyr 180
522 GATCGATACGGAACCAAGCAGCGGAAATATCTCCATCTCTCTGATATGAGAGTAT 581
181 AsnProGluValGlyTyrCysArgAspLysSerHisIleAlaLeuPheLeuLeuTyr 200
582 AACCCGAGAGTGGCTACGACAGGAGCTGAGCCACATGCGCCCTTGTCCTCTCAT 641
201 LeuProGluLysAspAlaPheTyrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220

642 CTTCCTGAGAGAGATCATCTGGGCACTGTGACAGCTGCTGGCCAGTGAAGGCACTCC 701
221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
702 CTGCGAGATTTTCACAGCCCAATGCGGAGCCGTCCAGGGGCTCCAGACCAACAGGAG 761
241 HisValAlaIleThrSerGlnProLysThrMetGlyHisGlnAspLysAspLeuCys 260
762 CATGTGATGCCACGTCACACCAACCAAGCCATCGGGCATCAGAGCAAGAAATCATCT 821
261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
822 GGGCAGATTTCCCGTTAGCTGCTCATCCGAAATGATGACCGGATCTCTCTCGGG 881
281 LeuThrLeuArgLeuThrAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
882 CTCACCTGCGCTGTGGAGCTGTATCTGTGAAGAGCAAGGCTTGATGCGGATA 941
301 ThrArgIleAlaPheLeuValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
942 ACAAGATGCTTTTAAGTTCAAGCAAGAGCCCTCAGCAAGACGTCAGTGTGCGCG 1001
321 ThrAlaArgPheCysAsnArgPheValAspThrTPAlaArgAspGluAspThrValLeu 340
1002 TGGGCACTTTTTCACACCGGTTCTGTGATCTGGGCAAGGAGTGAAGACACTGTGCTC 1061
341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
1062 AAGCATCTTAAGGCTCTATGAAGAACTAACAGAAAGCAGGGGACTTGCACCCCA 1121
361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgLysGlyLys 380
1122 GCCAAACCGACCAAGGCTGTCCGATCCAGGCTGTCCGGCTTCACTGTGGCGGAG 1181
381 ThrLeuCysLysGlyAspArgLysAlaProProGlyProProAlaArgPheProArgPro 400
1182 ACCCTCTGCAAGGGGAGCAGGAGCCCTCAGAGCCCAAGCCGCTTCCCGCGGCC 1241
401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
1242 ATTTGTCAGCTTCCCGCAGGAGCACTCTTCTTCCACACCTGTCTGTGGGCTC 1301
421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
1302 GTCCGGAAAGACACTTACCTTGTGGGCACTGAGGTGTCCAGCCGGCCCTGCTCAG 1361
441 GlyGlyProGlnGlySerTyrArgPheLeuGlnIleTyrAsnSerMetProArgLeuProThr 460
1362 GGAAGACCTCAGGATTCCTGAGATTCCTGAGTGAACCTCATGCGCCGCTCCCAAG 1421
461 AspLeuAspValGlnGlyProTyrPheArgHisTyrAspPheArgLysSerCysTyrVal 480
1422 GACTGTGAGAGGAGGCGCTTGTGGCTTGTGATTTAGACAGAGCTGTGGGCTC 1481
481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTyrGlnAlaGlnHisProAlaGlu 500
1482 CTGTCATATCCCAAGAGACCAAGCTGCGCCCTGTGTGAGGCTGAGACACCTTGGAG 1541
501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
1542 CGGGTGAATCGGCTTGTGCTGACCAAGCACTGATTCGACAGAGGACCCCTTCAAG 1601
521 AlaArgAspGlnGlnProCysAlaAlaProThrSerGlyProCysLeuCysGlyLeuHisIle 540
1602 GCTAGGAGCAAGACCGGTGTCTCCACCTCAGAGGCTTGTGCTGTGCGGCTTCACTTG 1661
541 GluSerSerGlnPheProProGlyLysPhe 549
1662 GAAAGTTCAGATTCCTTCCAGGCTTC 1688

RESULT 3
US-10-071-838-3
; Sequence 3, Application US/10071838

Publication No. US2003004481A1
GENERAL INFORMATION:
APPLICANT: Li, Jing
APPLICANT: Powers, Scott
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
FILE REFERENCE: 018781-007610US
CURRENT FILING DATE: 2002-05-07
PRIORITY FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
LENGTH: 2146
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
NAME/KEY: CDS
LOCATION: (1)..(1832)
US-10-071-838-3

Alignment Scores:

Pred. No.:	1,14e-258	Length:	2146
Score:	2907.50	Matches:	547
Percent Similarity:	89.67%	Conservative:	0
Best Local Similarity:	89.67%	Mismatches:	2
Query Match:	97.47%	Indels:	62
	14	Gaps:	1

US-10-071-838-2 (1-549) x US-10-071-838-3 (1-2146)

QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluAspIleIleMet 20
DB 1 ATGAGAGCTGTAGAGGTGCGGGCAGTTGTGGGCAAGAGCGAGAGCAATATTATG 60
QY 21 LysTrpGluLysGlyValAspAlaGlyLeuProGluAspIleProLysProPheArg 40
DB 61 AATATCGAAAGGAGCACCGAGCTGGCTCCAGAGCAAGAGGCGCTTAAAGCTTTTCA 120
QY 41 SerTrpAsnAspAspValAspHisLeuGlyTlleValHis----- 53
DB 121 AGCTACACACACACAGTGTGATCTTTGGGATTGTACAGTCTGCCGCTCTGGAGTCA 180
QY 53 ----- 53
DB 181 GCCCAGAGAGGCGCTTGTCTCCCTTCCCTTCTTCTGGGCTGAGCCCTGAG-240
QY 53 ----- 53
DB 241 CTGGAAGGAGCAGAGCCAGTCTTTCTGGGGGTGCGACCCAGGCTGGGGCCCTCAG 300
QY 54 ----- 54
DB 301 GCCCCTGAGTTCTCAGCTCGCCTGGGTTGCTTCACTGAGAGCGAGCTGCTCTCT 360
QY 60 LeuThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpVal 79
DB 361 CTGACTGGCGGGAGGCGAGCAAAATTCGGGGGAGTCAAGCCGAAAGACCAAGTGGT- 419
QY 80 AspMetLeuGlyAspTrpGluLysTrpLysSerSerArgLysLeuIleAspAlaGlyTyr 99
DB 420 GATATGCTGGGAGACTGGGAGAAATACAAAGACGAGAAAGCTCATAGATCAGAGGTAC 479
QY 100 LysGlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGlu 119
DB 480 AAGGAGATGCCCATGACATCCGGGGCCGATGTGTCACTCTCTGAAACATTAGAGAA 539
QY 120 MetLysLeuLysAsnProGlyArgTrpGlnIleMetLysLeuLysSerLysSerSer 139

DB 540 ATGAGGTGAAAGAAACCCCGAGAGATACAGATATCATAGAGAGGCAAGGCTCATCT 599
QY 140 GlnHisIleGlnHisIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePhe 159
DB 600 GAGCAGATCCAGGCGCATTCAGCCGGAGCTTAAGGGGACCTTAAGAGAGCATATATCTTC 659
QY 160 ArgAspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGlu 179
DB 660 AGGATCATAGAGAAACCAAGACGGGAACTACTTCACATCTCTCGGCATATAGAGAG 719
QY 180 TyrAsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaIleAspPheLeu 199
DB 720 TATTAACCCGAGGTGGCTACTGACAGGAGCTGAGGCCACATCGCGCTGTCTCTCTC 779
QY 200 TyrLeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHis 219
DB 780 TATCTTCTGAGAGAGATGATCTTGGGACTGTGAGCTGTGGCTGGCTAGAGAGCAC 839
QY 220 SerLeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGln 239
DB 840 TCCCTGCAAGGATTTCAAGCCCAATGGCGGAGCTCCAGGGGCTCCAGACCAACAG 899
QY 240 GlnHisValAlaIleThrSerGlnProLysThrMetGlyHisIleAspLysAspLeu 259
DB 900 GAGCATGTGTAGCAGTCAACCCAGACCATGGGCGCATCAGGACAAAGAAATCTA 959
QY 260 CysGlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyLysSerLeu 279
DB 960 TGTGGGAGAGTGTCCCGTTAGGCTGCTCATCCGGAATATTATGACGGATCTCTCTC 1019
QY 280 GlyLeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetPro 299
DB 1020 GGGCTACCTGCGCGCTGTGGAGCGGTATCTGTGAAGGAGCAAGCGCTTATATCCG 1079
QY 300 IleThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgCysGly 319
DB 1080 ATAAAGAGATGCCCTTTAAAGTTTCAAGAGAGCCCTCCAGAAAGTCCAGAGTGTGGC 1139
QY 320 ProTrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrVal 339
DB 1140 CCGTGGGACAGTTTGTGCAACCGGTCTGTATACCTGGGCGAGGATGAGACACTGTG 1199
QY 340 LeuLysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnLysAspLeuProPro 359
DB 1200 CTCAGACATCTTAGGCGCTTATAGAAAGAACTAACAGAAAGCAGGGGAGCTGCCACC 1259
QY 360 ProAlaLysProGluGlnLysSerSerAlaSerArgProValProAlaSerArgLysGly 379
DB 1260 CCAGCAAAACCCGAGCAAGGGTGTGTGCATCCAGGCTGTGGGCTTCACTGAGCGGG 1319
QY 380 LysThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArg 399
DB 1320 AAGACCTCTGCAAGGGGAGAGCAGAGCCCTCCAGGGCCACAGCCGGTTCCTGGGG 1379
QY 400 ProIleTrpSerAlaSerProProAlaProArgSerSerThrProCysProGlyGly 419
DB 1380 CCAATTTGTGACGTTCCCGGCAAGGCACTCTTCCACACCTCTGCTGTGGGG 1439
QY 420 AlaValArgLysAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeuAla 439
DB 1440 GCTGTCCGGGAAACACTTACCCTGTGGCACTCAGGGGTGGCCAGCCGGCTCTGGCT 1499
QY 440 GlnGlyGlyProGlnLysSerTrpArgPheLeuGlnTrpAsnMetProArgLeuPro 459
DB 1500 CAGGGAGACCTCAGGGTCTCTGAGATTCTGCAAGTGAACCTCAATGCCCTCCCTCCA 1559
QY 460 ThrAspLeuAspValGlnGlyProTrpPheArgHisLysTrpAspPheArgLysSerTrp 479
DB 1560 ACGACCTGAGAGTGAAGGCGCTTGTCTCGCATTAATATTACAGACAGAGCTGCTGG 1619
QY 480 ValArgAlaIleSerGlnLysAspGlnLeuAlaProCysTrpGlnAlaGlnHisProAla 499

Db 1620 GTCCTGTCATATCCAGAGAGACAGCTGGCCCCCTGCTGGAGGCTGAACACCTCGG 1679
Qy 500 GUAAGValArgSerAlaPheAlaAlaProSerThrAspSerArgGlnGlyThrProPhe 519
Db 1680 GAGCGGTGATGATCGGCTTTCCCTGTCACACGACATGATTCGACACGAGGACACCCCTTC 1739
Qy 520 ARGAlaArgAspGlnGlnProCysAlaProThrSerGlyProCysLeuGlyLeuHis 539
Db 1740 AAGAGTGAAGGACGACACGCGGTGCTCCACCTCAAGGCGCTTGCTTGCGGCTTCAC 1799
Qy 540 LeuGluSerSerGlnPheProGlyPhe 549
Db 1800 TTGGAAGTTCAGTTCCTCCAGGCTTC 1829
RESULT 4
US-10-302-172-692
Sequence 692, Application US/10302172
Publication No. US20040053250A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Dymnac, Radoje T.
TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids an
FILE REFERENCE: Polypeptides
FILE REFERENCE: 803.1CNC
CURRENT APPLICATION NUMBER: US/10/302,172
CURRENT FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 10/225,251
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: PCT US02/05095
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 09/799,451
PRIOR FILING DATE: 2001-03-05
NUMBER OF SEQ ID NOS: 950
SOFTWARE: pc FL_genes Version 2.0
SEQ ID NO 692
LENGTH: 2084
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (163) ..(1755)
US-10-302-172-692
Alignment Scores:
Pred. No.: 2,256-253 Length: 2084
Score: 2850.00 Matches: 528
Percent Similarity: 96.36 Conservative: 1
Best Local Similarity: 96.17 Mismatches: 2
Query Match: 95.54 Indels: 18
Gaps: 1
DB: 16
US-10-071-838-2 (1-549) x US-10-302-172-692 (1-2084)
Qy 1 MetAspValValGluValAlaGlySerTTPTPAlaGlnGluArgGluAspIleIleMet 20
Db 163 ARGAGGTGTAGAGGTCCGGGCACTGTGTGGCAACAGAGGAGGACATCATTAAG 222
Qy 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
Db 223 AATATGAAAGAAAGGACACCGAGCTGGCTGCCAGAGGACAGGCGCTTAAGCTTTTGA 282
Qy 41 SerTyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProPheLeu 60
Db 283 ACTATCAACACAACTCATCATCTTTGGGATTTGTACATGACGAGCTGCTCTCTCG 342
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTTPValAsp 80
Db 343 ACTGCGCGGAGCGGAGGAAATTCGGCGGAGATCAAGCGAAGAGCAAGTGGGTGAT 402
Qy 81 MetLeuGlyAspTTPGlyLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
Db 403 ATGCTGGAGAGACTGGGAGAAATACAAAGACAGCAAGAAAGCTCATATCATGAGGTACAG 462

Qy 101 GlyMetProMetAsnIleArgGlyProMetTTPSerValLeuLeuAsnIleGluGluMet 120
Db 463 GGAATGCCATGAAACATCCGGGGCCGAGTGTGTCTCTCTCAACCTAGAAATG 522
Qy 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerGlu 140
Db 523 AAGTTG----- 528
Qy 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
Db 529 AACATCCAGCGCATGACCGGACCTTAAGCGGACATTAGGAAACATATATCTTCAG 588
Qy 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluGly 180
Db 589 GATTCATAGGAAACAGACAGCGGAACTACTTCCATCTCTCTGCAATATAGAGAT 648
Qy 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
Db 649 AACCGGAGGTGGCTATCGAGGACCTGAGCCATCGCGCTTGCTCTCTCAT 708
Qy 201 LeuProGluGluAspAlaPheTTPAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
Db 709 CTTCCTGAGAGATGCAATTCGTGGCACTGTGCACTGCTGCGCAATGAGAGCACTCC 768
Qy 221 LeuGlnGlyPheHisSerProAsnGlnGlyThrValGlnGlyLeuGlnAspGlnGln 240
Db 769 CTGACGGGATTCACAGCCCAATAGCGGAGCGTCCAGGGCTCCACAGACCAACAGAG 828
Qy 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 829 CATGTGTATGCCACGTCAACACCAAGCAACAGGGGATCGACAGCAAGAAATCATGT 888
Qy 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyTlleserLeuGly 280
Db 889 GGGCAGGTTCCTCCCTATGAGCTGCTCATCCGATATGATTGACGGGATCTCTCGGG 948
Qy 281 LeuThrLeuArgLeuTTPAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
Db 949 CTCACCTCGGCTGTGGAGCTGTATCTGGTAGAAGAGGAAACAGCTTATGCCGAT 1008
Qy 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 1009 ACAAGATCGCTTTAAGTTACAGCAAGAGCGCTCAAGAAAGCTCCAGAGCTGGCCGG 1068
Qy 321 TTPAlaArgPheCysAsnArgPheValAspThrTTPAlaArgAspGluAspThrValLeu 340
Db 1069 TGGCACGTTTTTGCACCGGCTTCGTATACCTGGGCGAGGATGAGGACACTGTGCTC 1128
Qy 341 LysHisLeuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1129 AAGCATCTTAAGGCTCTATAGAACTTACAAAGAAAGCAAGGGGAGCTGCACCCCA 1188
Qy 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
Db 1189 GCCAAACCGAGCAAGGTCGTGGCATCCAGGCTGTGGCTTCAAGTGGCGGAAAG 1248
Qy 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1249 ACCCTCTGCAAGGGGAGACAGGAGCCCTCCAGGCCCAACAGCCGGTTCCCGGGCCC 1308
Qy 401 IleTTPSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyGlyAla 420
Db 1309 ATTTGTCAGCTTCCCGCACCGGACCTGCTTTCACACCTGTGCTGTGGGCT 1368
Qy 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1369 GTCGGGAAAGAACCTTATCTGTGGCACTCAAGGTGTGCCAAGCCGGCCCTGCTCAG 1428
Qy 441 GlyGlyProGlnGlySerTTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
Db 1429 GAGAGACCTCAAGGTTCTCTGAGATTCTGACAGTGAATCCATGCCCCGCTCCCAAG 1488

QY 461 AspleuAspValGluGlyProThrPheArgHisGlyTrpAspPheArgGlnSerCysTrpVal 480
 Db 1489 GACCTGACGTAGAGGGCCCTTGGTTCCGCAATATGATTCAGACAGAGCTTCGGGTC 1548
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCysTrpGlnAlaGlnIleProAlaGlu 500
 Db 1549 CCGTCCATATCCCAAGAGAGACAGACCTGGCCCTCGCTGGAGGCTGTAACACCCCTCGGAG 1608
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 Db 1609 CCGGTGAGATCGGCTTTCGCTGCACCCAGCACTGATTCGACAGAGGACCCCTTCAG 1668
 QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCysGlyLeuHisLeu 540
 Db 1669 GCTAGGACGACACGCGGTGTGCTCCACCTCAGGGCCCTTGCCTTCGCGCTCCACTTG 1728
 QY 541 GluSerGlnPheProProGlyPhe 549
 Db 1729 GAAAGTTCATGTTCCCTCCAGGCTTC 1755

RESULT 5
 US-10-302-172-691
 / Sequence 691, Application US/10302172
 / Publication No. US20040053250A1
 / GENERAL INFORMATION:
 / APPLICANT: Tang, Y. Tom
 / APPLICANT: Xue, Aidong J.
 / APPLICANT: Drmanac, Radoje T.
 / TITLE OF INVENTION: NO. US20040053250A1 Arginine-rich Protein-like Nucleic Acids an
 / FILE REFERENCE: 803 1CNCp
 / CURRENT APPLICATION NUMBER: US/10/302,172
 / PRIOR FILING DATE: 2002-11-21
 / PRIOR APPLICATION NUMBER: US 10/225,251
 / PRIOR FILING DATE: 2002-08-20
 / PRIOR APPLICATION NUMBER: PCT US02/05095
 / PRIOR FILING DATE: 2002-03-05
 / PRIOR APPLICATION NUMBER: US 09/799,451
 / NUMBER OF SEQ ID NOS: 950
 / SOFTWARE: pc_FL_genes Version 2.0
 / SEQ ID NO 691
 / LENGTH: 2072
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (163)..(1743)
 US-10-302-172-691

Alignment Scores:
 Pred. No.: 2,98e-251
 Score: 2827.00
 Percent Similarity: 95.63%
 Best Local Similarity: 95.63%
 Query Match: 94.77%
 Gaps: 16
 Indels: 22
 Matches: 2072
 Conservative: 0
 Mismatches: 2
 Gaps: 1

US-10-071-838-2 (1-549) x US-10-302-172-691 (1-2072)
 QY 1 MetAspValAlaGluValAlaGlySerTrpTrpAlaGlnGluArgGlnAspIleIleMet 20
 Db 163 ATGAGAGCTGTAGAGGTGCGCGGCGAGTGTGGCAACAGAGAGGAGCATCATATAG 222
 QY 21 LysGlyGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 Db 223 AAATACGAAAGGAGCACCGAGCTGGGCTGCAGAGGACAGAGGGGCTTAAGCTTTTGA 282
 QY 41 SerTrpAspAsnAsnValAspHisLeuGlyIleValHisGlyThrGluLeuProPheArg 60
 Db 283 AGCTACACACACACGTCGATCATTTTGGGATTTGTACATGACAGAGCTGCTCTCTG 342
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80

Db 343 ACTGGCGGAGGCGAAGCAAAATTCGGCGGGAATACGCCGAAAGGCAAGGGGTGAT 402
 QY 81 MetLeuGlyAspTrpGluLysGlySerSerArgLysLeuIleAspArgAlaGlyTrpLys 100
 Db 403 ATGCTGGGAGACTGGGAGAAATACAAAGACAGACAGAAAGCTCATATGATCGAGCTTCAAG 462
 QY 101 GlyMetProMetLeuIleArgGlyProMetCTPSPSerValLeuLeuAsnIleGlnGlnMet 120
 Db 463 GGAATGCCATGAAACATCCGGGCGCCGATGTGTGATCTCTTCAACACTGAGGAAATG 522
 QY 121 LysLeuLysAsnProGlyArgTrpGlnIleMetCysGluLysGlyLysAspSerSerGlu 140
 Db 523 AAGTTGAAAAACCCCGAGAGATACCAAGATCAATGAGAGAGAGGCAAGAGCTATCTGAG 582
 QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 583 CACATCCAGCGCATGACCGGAGCGTAAGCGGACATTAGAGAGCATATATCTTCAGG 642
 QY 161 AspArgTrpGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTrpGluGlnTrp 180
 Db 643 GATCGATACGGAACCAAGACGCGGAACTACTCACATCTCTCTGCAATATGAGAGATAT 702
 QY 181 AsnProGluValGlyTrpCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTrp 200
 Db 703 AACCGGAGAGTGGGCTACTCTCAGGAGACCTGAGCCACATCGCGCTTGTCTCTCTAT 762
 QY 201 LeuProGluGluAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 763 CTCTCTAGAGGATCATCTCGGCGACGTGTCAGCTGCGGCACTGAGAGCACTCC 822
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
 Db 823 CTCGAGGAGATTCACAGCCCAATATGCGGACCTCCAGGCGGCTCCCAAGCCAAAGAG 882
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 883 CATGTGTAGCCACGTCAACCAAGCAGCTGGGGCATAG----- 924
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 925 -----ATCTCTCCGG 936
 QY 281 LeuThrLeuArgLeuTrpAspValTrpLeuValGluGluGlnAlaLeuMetProIle 300
 Db 937 CTCACCTCGGCGCTGTGGAGCGTATCTGTAGAGAGGAGCAAGCGTATATCCCATATA 996
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 997 ACAGAAATCCCTTTAAGGTTCCAGAGAGAGCCCTCAGAGAGAGAGTCCAGGTGAGCCCG 1056
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspTrpValLeu 340
 Db 1057 TGGGCAAGTTTTCGACACCGGTTGTTGATCTCGGCGCAGAGTACAGACATCTGTCTC 1116
 QY 341 LysHisLeuAspAlaSerMetLysLysLeuThrArgLysGlnIleLysLeuProPro 360
 Db 1117 AACCATCTTAAGGCGCTTATAGAAAGAACTAACAGAAAGAGGGAGACCTGCAGACCCCA 1176
 QY 361 AlaLysProGluGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLysLys 380
 Db 1177 GCCAAACCCGAGCAAGGAGTGTGCGATCCAGGCGCTGTCCGGCTTCACTGCGGAGAG 1236
 QY 381 ThrLeuCysLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 1237 ACCCTTGCAGAGGAGGAGCAGGAGGCGCTCCAGGCGCCACAGCCGGGCTCCCGGCGCC 1296
 QY 401 IleTrpSerAlaSerProProArgAlaProArgSerThrProCysProGlyGlyValAla 420
 Db 1297 ATTTGTGACACTTCCCGGCAAGGAGACCTGTCTTCCACACCTGCTCCGAGGGGCT 1356
 QY 421 ValArgGluAspThrTrpProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440

Db 1357 GTCGGGAAGACCTACCTGTGGCACTACAGGTGTGCCAGCCCGCTGCTCAG 1416
Qy 441 GIVGLYPRGNGLYSEITPARGPHELEUNGINTPANSERMEPROARGLEUPROTHR 460
Db 1417 GAGAGACCTCAGGGTTCGAGAGATTCCTGCACTGAGAACTCCAGCCGCTCCACAAG 1476
Qy 461 AASPLeuASPValGluGlyProTTPHearGHIATYAspPheARGINSerCYSTRVal 480
Db 1477 GACCTGAGAGTGAAGGGCCCTTGCTCCGCAATTATGATTCAGACAGAGCTGCTGGTTC 1536
Qy 481 ARGVALIISERGLINGLUAPGLINLEUALAPROCYETFTGGINAAGLUHISPROAAGLU 500
Db 1537 CDTGCAATATCCAGAGAGACAGCTGCGCCCTGCTGGCAGGCTAACAACCTGCGAG 1596
Qy 501 ARGVALIISERGLINLEUALAPROSERTHASPSERASPGINGLYTHPROPHARG 520
Db 1597 CCGGAGAGATCGGCTTCCGTGCAACCAAGCATGATTCGACAGGAGACCCCTTCACA 1656
Qy 521 ALARGASPGIUGINPROCYBALAPROTHSERGLYPROCYLEUCYSGLYLEUHIISLEU 540
Db 1657 GCTAGGAGAGAACAGCCGTGTGCTCCCACTCAGGGCCCTTGCTGCGGCCCTCCACTTG 1716
Qy 541 GLUSERSEGLINPHEPROPGLYPHE 549
Db 1717 GAAAGTTCTCAGTTCCTCCAGGCTTC 1743

RESULT 6
US-10-108-260A-676
/ Sequence 676, Application US/10108260A
/ Publication No. US20040005560A1
/ GENERAL INFORMATION:
/ APPLICANT: HELIX RESEARCH INSTITUTE
/ TITLE OF INVENTION: No. US20040005560A1el full length cDNA
/ FILE REFERENCE: H1-A0106
/ CURRENT APPLICATION NUMBER: US/10/108,260A
/ NUMBER OF SEQ ID NOS: 5458
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 676
/ LENGTH: 2647
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-108-260A-676

Alignment Scores:
Pred. No.: 4,016-251 Length: 2647
Score: 2827.00 Matches: 525
Percent Similarity: 95.63% Conservative: 0
Best Local Similarity: 95.63% Mismatches: 2
Query Match: 94.77% Indels: 22
DB: 16 Gaps: 1

US-10-071-838-2 (1-549) x US-10-108-260A-676 (1-2647)
Qy 1 MetAspValIaIGluValaIaGlySEITPTPAaIaGInGluArgGluAspIleIleMet 20
Db 755 ATGCACTGTGTAGAGTCCGCGGCACTGTGTGGGCAACAAGGAGAGACATCATTAAG 814
Qy 21 LysTYRGluLysGLYHISARGaIaGlyLeuProGluAspLysGLYProLysProPheArg 40
Db 815 AATAACGAAAGGACACCGAGCTGGGCTGCCAGAGACAAAGGGGCTTAAGCCTTTTCA 874
Qy 41 SEITTPAAspAspValaIaSPHISLEUNGlyIleValHISGLuThrGluLeuProProLeu 60
Db 875 ACCTACAAACAAACCTCATCTTTGGGAGATTGTACATGAGACGAGCTGCTCTCTG 934
Qy 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerIYSTRValaAsp 80
Db 935 ACTGCCCGGAGAGCGAGACAAATTCGCGGAGATCAGCCGAAAGACAGAGTGGTGAT 994
Qy 81 MetLeuGlyAspTPRGluLysTYRlySerserArgLysleuIleAspArgAlaTYRlys 100
Db 995 ATGCTGGAGACTGGAGAAATACAAAGACAGCAAGAACTCATAGATGAGCGTACAG 1054

Qy 101 GYMeTPrometAsnIleArgGlyProMetTTPSERValleuLeuAsnIleGluGluMet 120
Db 1055 GGAATCCCATBAACATCCGGGGCCCGATGTGTGATCTCTCTGAACACTAGAGAAAG 1114
Qy 121 LysLeuLysAspProGluYArgTYRGlnIleMetLysGluLysGLYLYAspSerserGlu 140
Db 1115 AAGTTGAAAAACCCCGAAGATACCAAGATCATGAAGAGAAAGGCAAGAGCTCATCTGAG 1174
Qy 141 HISIleGlnArgTlaAspArgAspValSerGlyThrLeuArgLysHISIlePhePheArg 160
Db 1175 CACATTCAGCGCATGACCGGAGACGTAAAGCGGAGACATTAAAGAGCATATATTCTTCAG 1234
Qy 161 AspArgTYRGlyThrLysGlnArgGluLeuLeuHISIleLeuLeuAlaTYRGLuGlyTYR 180
Db 1235 GATCATACGGAACCAAGAGCGGGAACATCTCCACATCTCTCTGGCATATAGAGATAT 1294
Qy 181 AsnProGluValGlyTYRcyAspArgAspLeuSerHISIleAlaIleuPheLeuLeuTYR 200
Db 1295 AACCGGAGAGTGGGCTACTGACAGGAGCTGAGACCATGCGCCCTTGCTCTCTCAT 1354
Qy 201 LeuProGluGluAspAlaPheThrPalaLeuValGlnLeuLeuAlaSerGluArgHISer 220
Db 1355 CTTCCTGAGAGAGATGATTCATTTGGGCACTGTGACGTGCTGCGCAAGAGGACATTC 1414
Qy 221 LeuGlnGlyPheHISerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 1415 CTGCAAGGATTTCAAGCCCAATGCGGAGCCGCTCCAGGGCTCCMAAGCAAGAG 1474
Qy 241 HISValIaIaIaThSerGlnProLysThMetGlyHISGlnAspLysAspLeuCYs 260
Db 1475 CATGTGTGTGACAGCTCAACCAACCAACCATGGGGCATCAG----- 1516
Qy 261 GLYGlnCYserProLeuGlyCYleuIleArgIleleuIleAspGlyIleSerLeuGly 280
Db 1517 -----ATCTCTCTCGGG 1528
Qy 281 LeuThrLeuArgLeuThrAspValTYRleuValGlnGlyGluAlaLeuMetProIle 300
Db 1529 CTCACCTCGGCTGTGGACCTGTATCTGTGAGAAAGGAGAACAGCGTTGATGCCGANA 1588
Qy 301 ThrArgIleAlaPheLysValaGlnGlnLysArgLeuThrLysThSerArgCYsglyPro 320
Db 1589 ACAAGATCGCTTTAAGTTCAAGCAAGAGCCCTCAACAGACATCGTCAGTGGCCCG 1648
Qy 321 TPAlaArgPheCYAspAspArgPheValAspThrTPAlaArgAspGluAspThrValLeu 340
Db 1649 TGGGACAGTTTTTGCAACCGGTTCTGTATACCTGGGCAAGGAGATGAGACACTGTGCTC 1708
Qy 341 LysHISleuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
Db 1709 AAGCATCTTAGGGCTCTATGAGAAACTACAAAGAAAGAGGGGAGCTGCACCCCA 1768
Qy 361 AlaLysProGluGlnGlySerserAlaSerArgProValProAlaSerArgGlyGlyLys 380
Db 1769 GCCAAACCGGACAAAGGATCGTCGCAATCCAGCGCTGCGGCTTCAAGTGGCGGAGAG 1828
Qy 381 ThrLeuCYLysGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
Db 1829 ACCCTCTGAGGGGAGACAGGAGCCCTCCAGGCCCAACAGCCCGGTTCCCGGGGCC 1888
Qy 401 IletPserAlaSerProProArgAlaProArgSerserThProCYserProGlyGlyAla 420
Db 1889 ATTTGATGACTTCCCGGACAGGCACTGCTTTTCCACACCTGTCTGTGTGGGCT 1948
Qy 421 ValArgGluAspThrTYRProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
Db 1949 GTCGGGAAAGACACTACCTGTGGGCACTGAGGGTGTGCCAGCCGCGCTGCTCAG 2008
Qy 441 GIVGLYPRGNGLYSEITPARGPHELEUNGINTPANSERMEPROARGLEUPROTHR 460
Db 2009 GAGAGACCTCAGAGGTTCTTGAGATTCTGCAAGTGAACCTCATGCCCCGCTCCCAAG 2068


```
QY 461 AspleuaspVal1GluGlyProTrrPheargHisTyrAspPheargGlnSerCystrVal 480
Db 2069 GACCTGGACGTAAGAGGCGCTTGGTTCGGCATTATGATTGACAGAGGCTGCTGTC 2128
QY 481 ArgAla1IleSerGlnGluAspGlnLeuAlaProCystrPrlna1agiun1aspAlaGlu 500
Db 2129 CGTCCATATCCCGAGGAGACCGAGCTGCGCCCTGGCTGGAGGCTGAACCCCTGGAG 2188
QY 501 ArgValArgSerAlaPheAla1AproSerThrAspSerAspGlnGlyThrProPhearg 520
Db 2189 CGGGTAGATCGGCTTTCGCTGCAACCGACATGATTCGACAGGAGCCCTTCAG 2248
QY 521 AlaArgAspGluGlnProCyAlaProThrSerGlyProCyLeuGlyLeuHis1Ileu 540
Db 2249 GCTAGGACGACAAAGCGGTGTGCTCCACCTGAGGCGCTTGCCTTCGGCGCTTCCACTTG 2308
QY 541 GluSerSerGlnPheProGlyPhe 549
Db 2309 GAAAGTTCTCAGTTCCTCCAGGCTTC 2335

RESULT 7
US-10-094-466-61
/ Sequence 61, Application US/10094466
/ Publication No. US2003020363A1
/ GENERAL INFORMATION:
/ APPLICANT: Spylek et al.
/ TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
/ TITLE OF INVENTION: AND METHODS OF USING
/ FILE REFERENCE: THE SAME
/ CURRENT APPLICATION NUMBER: US/10/094,466
/ PRIOR APPLICATION NUMBER: 2002-03-07
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/274,281
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/288,148
/ PRIOR FILING DATE: 2001-05-02
/ PRIOR APPLICATION NUMBER: 60/274,849
/ PRIOR FILING DATE: 2001-03-09
/ PRIOR APPLICATION NUMBER: 60/275,235
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: 60/338,375
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 60/275,579
/ PRIOR FILING DATE: 2001-03-13
/ PRIOR APPLICATION NUMBER: 60/335,302
/ PRIOR FILING DATE: 2001-10-31
/ PRIOR APPLICATION NUMBER: 60/275,601
/ PRIOR FILING DATE: 2001-03-13
/ PRIOR APPLICATION NUMBER: 60/276,000
/ PRIOR FILING DATE: 2001-03-14
/ PRIOR APPLICATION NUMBER: 60/277,338
/ PRIOR FILING DATE: 2001-03-20
/ PRIOR Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 114
/ SOFTWARE: Patin 2.1
/ SEQ ID NO 61
/ LENGTH: 1752
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (31)..(1678)
US-10-094-466-61

Alignment Scores:
Pred. No.: 2,296-247 Length: 1752
Score: 2784.00 Matches: 515
Percent Similarity: 96.17% Conservative: 13
Best Local Similarity: 93.81% Mismatches: 21
Query Match: 93.33% Indels: 0
DB: 15 Gaps: 0

US-10-071-838-2 (1-549) * US-10-094-466-61 (1-1752)
```

```
QY 1 MetAspVal1a1GluVal1a1agiysertPrla1agiun1aarg1uasp1IleMet 20
Db 31 ATGACCTGGTAAAGGTCCGGGTAGTGTGGTGGCAAGAGCGAGACATCTTATG 90
QY 21 LysTyrGlyu1ags1YH1sArgAlaGlyLeuProGluAspLysGlyProLysProPhearg 40
Db 91 AAATACGAAGAAAGGACACCGAGCTGGCTGGCCAGAGACAAAGGGCTTAAGTCTTTGCA 150
QY 41 SerTyrAspAspAsnVala1sphi1sleuGly1IleVal1H1sGluThrGluLeuProLeu 60
Db 151 AGCTACAAACAAACAGCTCATTTGGGAGATTGTAACAGAGACGAGACTCTCTCTG 210
QY 61 ThrAlaArgGluAla1LysGln1IleArgGln1IleSerArgLysSerLysTyrVala1asp 80
Db 211 ACTGGCGGGAGAGTGAAGCAATTCGGGGAGATGACCCAAAGACCAAGTGGGTGAAA 270
QY 81 MetLeuGlyAspTrrPrlu1LysTyrLysSerSerArgLysLeu1IleAspArgAla1TyrLys 100
Db 271 ATGCTGGGAGATGGAGACCTACAAAAACAGCAAGAAAGCTCATGATTCAGGCTACAG 330
QY 101 GlyMetProMetAsn1IleArgLysProMetTrrSerVal1LeuLeuAsn1IleGluMet 120
Db 331 GGAATTCCTCATGAACATCCGGGGCCCATGTGTGCTCTCTGAAACATTGAGAAATC 390
QY 121 LysLeuLysAsnProGlyArgTyrGln1IleMetLysGluLysGlyLysArgSerSerGly 140
Db 391 AAGTTGAAAAACCCCGGAAGATACCATCATGTAAGAGAAAGGCAAGGCTATCTGAA 450
QY 141 His1IleGlnArg1IleAspAspVala1SerGlyThrLeuArgLysHis1IlePhePhearg 160
Db 451 CACATCCAGCAGATGAGCTGAGCGTAAGCGGAGACATTAAGAGGACATATATCTTCAAG 510
QY 161 AspArgTyrGlyThrLysGlnArgLysLeuLeuHis1IleLeuLeu1a1ArgGluTyr 180
Db 511 GATGATACGGAACCAAGACGCGGAACCTATTACCTCTCTGCGATGAGAGAT 570
QY 181 AspProGluVal1GlyTyrCyAspAspLysSerHis1IleAla1IlePheLeuLeuTyr 200
Db 571 AACCCGAGAGTGGCTACCTGACGAGACCTGAGCCATGCGCGCTTCTCTCTTAT 630
QY 201 LeuProGluGluAspAlaPheTrrAlaLeuVala1GlnLeuLeuAlaSerGluArgHis1Ser 220
Db 631 CTTCCTGAGAGGATCATTTCTGGGACATGTCGACCTCTGCGACATGAGAGGACCTCC 690
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrVala1GlnGlyLeuGlnAspGlnGlnGlu 240
Db 691 CTCAGGGATTTCACAGCCCAATGCGGACCGTCCAGGGGCTTCAAGACCAAGAG 750
QY 241 HisVala1Ala1ThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
Db 751 CATGTGTAGCCACGTCATACCCAAACACATGTGCGATCGACAAAGAAAGATCTATGT 810
QY 261 GlyGlnCySerSerProLeuGlyCysLeu1IleArg1IleLeu1IleAspGly1IleSerLeuGly 280
Db 811 GGGCAGTGTTCCTCTTGGCTGCTGCTCATTCGGAATTTGACGGGATCTCTCGGG 870
QY 281 LeuThrLeuArgLeuTrrAspVala1TyrLeuVala1GlnGlyGlnGlnAlaLeuMetPro1Ile 300
Db 871 CTCACCTTGGCGGCTGTGGAGCTGTATCTGTTAAAGCAAGGCGGTTGAATGCCGANA 930
QY 301 ThrArg1IleAlaPheLysVala1GlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
Db 931 ACNAGAAATGCTTTAAGTTCACTGACAGCGCTCCGAAAGCTCCAGGTGTGCGCG 990
QY 321 TrrAlaArgPheCysAsnArgPheVala1AspThrTrrAlaArgAspGluAspThrVala1Leu 340
Db 991 TGGGACGCTTTTGGAAACCGTTCGTGATGCTTGGGCCAGGAGATATACACTGTGCTC 1050
QY 341 LysHis1IleuArgAla1AspMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
Db 1051 AAGCATCTTGGGCTCTATGAAGAAACTTAACAAAGAAAGGAGGAGCTGCGACCCCA 1110
```


QY 361 AAlATyPProGluGlnGlySerSerAlaSerAProAlaProAlaSerAglYglYlys 380
 DB 1111 GCCAAACCCGACGAGGCGTCCGCGATCCAGCCCTGCGACGTTCACTGCGCGGAG 1170
 QY 381 ThrLeuCylyyGlyyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 DB 1171 ACCCTCTGCAAGGGGAGCAGGAGGCCCTCCAGGCCACAGCCGGGTTCCCATGCGCC 1230
 QY 401 IlleTPSerAlaSerProProAlaArgAlaProArgSerSerThrProCyAspProGlyAla 420
 DB 1231 ATTGGTCAGCTTCCCGCCAGCGGACCTCGTTCCACACCCCTGTCCTGCGGCGCT 1290
 QY 421 ValAspGluAspThrTyPProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 DB 1291 GTCCGGAGAGACCTTACCTCTGGGCACTCAGGGTGTCCAGGCCCGGCTGCTCAG 1350
 QY 441 GlyGlyProGlnGlySerTPArgPheLeuGlnTPAsnSerMetProArgLeuProThr 460
 DB 1351 GGAGGACCTCAGGGTTCCTGAGATTCTCGAGTGGAATCCATGCCCCGCTCCCAAG 1410
 QY 461 AspLeuAspValGluGlyProTPArgPheArgHisTyTPAspPheArgGlnSerCySTPVal 480
 DB 1411 GACCTGGACGTAGGGGACCCCTTGTTCCGCGTATGATTTGACAGAGCTGTGGGTC 1470
 QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCySTPArgAlaGlnIleAspAlaGlu 500
 DB 1471 CTTGCCATATCCAGAGAGACAGCCGCGCACCTGCTGGCAGGCTGAGACACCTTGGAG 1530
 QY 501 ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
 DB 1531 CGGGTGAATCGGGCTTTTCAGTCCGCGCCAGACATGATTCGACAGAGGACCCCTTCAGA 1590
 QY 521 AlaAspArgGluGlnProCyAlaAlaProThrSerGlyProCyAlaLeuCyGlyLeuHisLeu 540
 DB 1591 GCTAGGAGCAGAACACAGTGTCTCCACTCAGACCTTGCTGTGGCGCTCCACTTGG 1650
 QY 541 GluSerSerGlnPheProProGlyPhe 549
 DB 1651 GAAAGTTCTCAGTTCCTCCAGGCTTC 1677
 RESULT 8
 US-10-071-838-5
 ; Sequence 5, Application US/10071838
 ; Publication No. US20030044814A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Li, Jing
 ; APPLICANT: Powers, Scott
 ; APPLICANT: Xiang, Phil
 ; APPLICANT: Peng, Yue
 ; APPLICANT: Tularik Inc.
 ; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
 ; FILE REFERENCE: 018781-007610US
 ; CURRENT APPLICATION NUMBER: US/10/071,838
 ; PRIOR FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: US 60/267,615
 ; PRIOR FILING DATE: 2001-02-08
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 5
 ; LENGTH: 1862
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1538)
 US-10-071-838-5
 Alignment Scores:
 Pred. No.: 2,566-246 Length: 1862
 Score: 2773.00 Matches: 515
 Percent Similarity: 93.81% Conservative: 0

Best Local Similarity: 93.81% Mismatches: 0
 Query Match: 92.96% Indels: 34
 DB: 14 Gaps: 1
 US-10-071-838-2 (1-549) x US-10-071-838-5 (1-1862)
 QY 1 MetAspValAlaGluValAlaGlySerTPArgPheArgGlnAlaArgGlnIleMet 20
 DB 1 ATGACGTGTAGAGTGGCGGCGAGTTGTGGGACAGAGCGAGAGCATCTATTG 60
 QY 21 LysTyGluYyGlyyHisArgAlaGlyLeuProGluAspGlyProLysProPheArg 40
 DB 61 AAATGCAAAAGGACACCGAGTGGCTGGCAGAGGCAAGGCGCTTAAGCTTTTGA 120
 QY 41 SerTPAsnAspValAspHisLeuGlyIleValHisGluThrGluLeuProLeu 60
 DB 121 AGCTACAAACACACGTGATCATTTGGAGATTGACATGAGACGAGCTGCTCTG 180
 QY 61 ThrAlaArgGluAlaGlyGlnIleArgArgGluIleSerArgLysSerLysTPArgAla 80
 DB 181 ACTGCGCGGAGGCGAAGCAAAATTCGCGGAGATCACCCGAAAGACAGATGGTGAT 240
 QY 81 MetLeuGlyAspTPGluLysTyLysSerSerArgLysLeuIleAspArgAlaTyLys 100
 DB 241 ATGCTGGAGACTGGGAGAAATACAAAGACAGAAAGCTCATGATGAGCGTACAG 300
 QY 101 GlyMetProMetAsnIleArgGlyProMetTPSerValLeuLeuAsnIleGluGlnMet 120
 DB 301 GGAATGCCCATGAAATCCGGGCGCCGATGTGTAGTCTCTGTAACATGAGGAAATG 360
 QY 121 LysLeuLysAsnProGlyArgTyGlnIleMetLysGluYyGlyLysArgSerSerGlu 140
 DB 361 AAGTTGAAAAACCCCGAAGATACCATGATCAGAGAGAGGAGGACAGTCACTAG 420
 QY 141 HisIleGlnArgGlyAspArgAspValSerGlyThrLeuArgGlyHisIlePhePheArg 160
 DB 421 CACATCCAGCCCATGACCGGGACCTTAAGCGGACATTAAAGAACATATATCTTCAGG 480
 QY 161 AspArgTyArgLysThrLysGlnArgGluLeuHisIleLeuLeuAlaTyGluGluTy 180
 DB 481 GATCGATAGGAAACACAGCAGCGGAGACTACCTCCACTCTCTGCAATAGAGAGAT 540
 QY 181 AsnProGluValAlaGlyTyCyArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTy 200
 DB 541 AACCGGAGGTGGCTACTGCAAGGACCTGAGCACAATCGCGCTTCTCTCTCAT 600
 QY 201 LeuProGluGluAspAlaPheTPArgAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 DB 601 CTTCCTGAGAGAGATGATTCGGGCACTGGTGCACTGTGGCCAGAGAGGCACTCC 660
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
 DB 661 CTGCAAGGATTTCAACAGCCCAATGGCGGACCGTCCAGGGGCTTCAAGACCAAGAG 720
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspGlyLysAspLeuCy 260
 DB 721 CATGTGTAGCCAGTCAACACCCAGACATGGGGCATCAG----- 762
 QY 261 GlyGlnCySerProLeuGlyCyLeuLeuIleArgIleLeuIleAspGlyLysSerLeuGly 280
 DB 762 ----- 762
 QY 281 LeuThrLeuArgLeuThrAspValTyTrLeuValGluGluGlnAlaLeuMetProIle 300
 DB 763 -----TATCTGTAGAAAGCGAAGACAGCGCTTGAATGCCATA 798
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCyGlyPro 320
 DB 799 ACAAAATCGCTTTAAGTTCAAGAGAGAGCGCTCAAGAGACGTTCAAGGTGTGGCCG 858
 QY 321 TrpAlaArgPheCyAsnAspArgPheValAspThrTPArgAlaArgAspGluAspThrValLeu 340
 DB 859 TGGGCACTTTTGGAAACCGGTTGTTGATACCTGTGGCCAGAGGATGAGAGACATGTGCTC 918

QY 341 LysHsleuArgAlaSerMetLysLysLeuThrArgLysGlnGlyValAspLeuProPro 360
 Db 2714 MAGCATCTTAGGGCCCTCTACGAAGAACTACAGAGAACAGAGGAGCCTGCCACCCCA 2773
 QY 361 AAlAsProGlnGlnGlySerSerAlaSerArgProValProAlaSerArgGlyLys 380
 Db 2774 GCCAAACCGGACCAAGGGCTCTTGGGACCCAGGCTGTGCCGCTTCACTGGTGGAG 2833
 QY 381 ThrLeuCyLeuGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
 Db 2834 ACCCTCTGACAGGGGATATAGGAGGCGCTCTCAAGGCCACAGCCCAAGTTCAGCGGCC 2893
 QY 401 IleTSPSerAlaSerProProAlaPheArgSerSerThrProCySerProGlyAla 420
 Db 2894 ATTTCCTAGCTTCCCGCCGACAGGCGATCTCGTTTCCAGCCCTGTCTGTGGGGCT 2953
 QY 421 ValArgLysPheThrTyProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
 Db 2954 GTCCGGAGAACACGATACCTGTGGGCACTCAGGGGTGCCACGCTGGCCCTGGCTCAG 3013
 QY 441 GlyGlyProGlnGlySerThrPheLeuGlnThrPheSerMetProArgLeuProThr 460
 Db 3014 GGAGAGCCTCAGAGGTCTCTGAGATTCCTGAGGTGAGTCAATGCCCGGCTCCCAAG 3073
 QY 461 AspLeuAspValGlyGlyProThrPheArgHsIleTyAspPheArgGlnSerCySerPVal 480
 Db 3074 GACCTGGATATAGGGGGCCCTTGGTTCCTCCCATATGATTTTGAACGAGCTGTGGTGC 3133
 QY 481 ArgAlaIleSerGlnGlyAspGlnLeuAlaProCySerGlnAlaGlnHsIleProAlaGln 500
 Db 3134 CGTGCATATCCCAAGAGGACCAAGCTGCGACCTGTGCAAGCTGACATCGCGGAGAG 3193

RESULT 10

US-10-283-975A-272
 ; Sequence 272, Application US/10283975A
 ; Publication No. US20040110792A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ortho-Clinical Diagnostics, Inc.
 ; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
 ; FILE REFERENCE: CDS 293 PCT
 ; CURRENT APPLICATION NUMBER: US/10/283,975A
 ; PRIOR FILING DATE: 2002-10-30
 ; PRIOR APPLICATION NUMBER: 60/340,938
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/338,997
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/340,081
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/341,012
 ; NUMBER OF SEQ ID NOS: 900
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 272
 ; LENGTH: 8201
 ; TYPE: DNA
 ; ORGANISM: HUMAN
 US-10-283-975A-272

Alignment Scores:

Pred. No.: 1,85e-186 Length: 8201
 Score: 2133.50 Matches: 408
 Percent Similarity: 77.23% Conservative: 33
 Best Local Similarity: 71.45% Mismatches: 58
 Query Match: 71.52% Indels: 72
 DB: 17 Gaps: 2

US-10-071-838-2 (1-549) x US-10-283-975A-272 (1-8201)

QY 1 MetAspValValGluValAlaGlySerTTPRAlaGlnGluArgGluAspIleIleMet 20
 Db 1697 ATGGACATGGTATGAGAAATGACATATGTTTGCAGGSCACAGAGCGGAACATATCTATG 1756
 QY 21 LysTyGlnLysGlyHsIleArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40

Db 1757 AAGTATGACAAAGGACACCGAGCTGGGCTCCAGAGAGCAAGGGGCTAGCCCTT 1813
 QY 41 SerTyAsnAsnAsnValAspHsIleuGlyIleValHsIleThrGlnLeuProProLeu 60
 Db 1814 GGAATCAACAGACAGATGATGTTTGGCATTTTGTGATGAGACGAGCTGCTCTGG 1873
 QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerTyTrpValAsp 80
 Db 1874 ACTGCACGGAGGCGAAGAAATTCGGCGGAGATGACAGAACAGCAAGTGGATGAA 1933
 QY 81 MetLeuGlyAspTrpGlnLysTyLysSerSerArgLysLeuIleAspArgAlaTyTrpLys 100
 Db 1934 ATGCTGGAGATGAGAGACATATAGCACTAGCAAACTCATGATGATGATGATGAA 1993
 QY 101 GlyMetProMetAsnIleArgGlyProMetTSPSerValLeuLeuAsnIleIleuIleMet 120
 Db 1994 GGAAATCCCATMAATCCGGGGCCCGGTGTGTGATCTCTTCAACATTCAGAAATC 2053
 QY 121 LysLeuLysAsnProGlyValArgTyGlnIleMetLysGlnLysGlyLysArgSerGln 140
 Db 2054 AAGTTGAATAACCCGGAAGATACAGATCATGAAGAGAGGCGAAGGTCACTGAA 2113
 QY 141 HsIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHsIlePhePheArg 160
 Db 2114 CACATCCACCAATGACCTGAGCCTGAGAGAGACGATCTCCGGAACCATGCTCTTTAG 2173
 QY 161 AspArgTyGlyTyThrLysGlnArgLysLeuLeuHsIleLeuLeuAlaTyArgLysTy 180
 Db 2174 GATCGATATGAGCGCAAGAGAGGAGAACTATTCATCTCTCTGCGCATTCGAGAT 2233
 QY 181 AsnProGluValGlyTyCyAspArgAspLeuSerHsIleAlaLeuPheLeuTy 200
 Db 2234 AACCCGAGGTGGCTACTGACGAGACCTGAGCCATCAACGCCCTTGTCTCTTAT 2293
 QY 201 LeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGlnArgHsIle 220
 Db 2294 CTGCTGAGAGAGAGACGATCTGGGACGAGTGGAGCTGCTGCCAGTGAAGGACCTC 2353
 QY 221 LeuGlnGlyPheHsIleSerProAsnGlyGlyTyThrValGlnGlyLeuGlnAspGlnGln 240
 Db 2354 CTGCCAGATTCACAGGCCCAATGTTGGAGACGTCCAGGGGCTCCAGAACCAAGAG 2413
 QY 241 HsValValAlaThrSerGlnProLysThrMetGlyHsIleAspLysLysAspLeuCy 260
 Db 2414 CATGTGTACCCCAAGTCAACACCAAGACATGTGCATTCGAGACAGGAAGCTTATGC 2473
 QY 261 GlyGlnCySerProLeuGlyCySerLeuLeuArgIleLeuIleAspGlyIleSerLeuGly 280
 Db 2474 GGGCAGTGTGCTGCTGTTAGGCTGCTCTCCGGAACGTGATTGAGGGATCTCTCGGG 2533
 QY 281 LeuThrLeuArgLeuTrpAspValTyLeuValGlnGlyGlnAlaLeuMetProIle 300
 Db 2534 CTCACCCCTGCGCTGTGGACCTGTATTGTGTGAAGAGAAACAGATGTGATGCCATA 2593
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysSerArgGlyPro 320
 Db 2594 ACCAGCATTTGCTTAAGGTTTACGAGAGAGGCTCTCATGAAACATCCAGGTGGCTGG 2653
 QY 321 TrpAlaArgPheCyAsnAspPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 2654 TGGGACGTCTCGGAGAACCAATCTTCGATACCTGGGCAATGAACAGTACACCGTGC 2713
 QY 341 LysHsIleuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProPro 360
 Db 2714 AAGCATTTAGGGCTCTACGAAGAACTAACAGAAAGCAAGGGGACCTGGACCCCA 2773
 QY 360 ----- 360
 Db 2774 GGCCCAACAGCCCTGGAGCAAGGTGTGTGCAAGAGCCCGACCACTGAAACCTGG 2833
 QY 360 ----- 360

Db 2834 GGGCAGTCCAGAGACCAACCAATGCCCCAAGCGCTTCCCATGCGAGGACAC 2893
QY 360 ----- 360
Db 2894 ACCCTCCCTTGGATCAGCAGACTACAGCGCTGTCTCAGTGTACAGACCAAGGAGCC 2953
QY 361 ----- 361
Db 2954 ACACAGAGACCCCAAGAGACTCCAGAGATGACCGCAACCGCAAGAGGCTTGGCACC 3013
QY 370 rARgProValProAlaSerArgIyGlyThLeuCyAlaGlyAspArgIyAlaPr 390
Db 3014 CAGGCTGTGCGGCTTACGTGTGTGGAGAACCTCTGCAGAGGGATAGAGAGCC 3073
QY 390 OProGlyProProAlaArgPheProArgProIleTpsSerAlaSerProProArgAlaPr 410
Db 3074 TCCAGGCCACACACCCAGTTCAGCGGCCCATTTGCTACGCTTCCCGCCCATGGGCATC 3133
QY 410 OArgSerSerThrProCysProGlyGlyAlaValAlaArgIyAspThrTyProValGlyTh 430
Db 3134 TCGTTTTCACGCGCTCTCTGTGTGGGCTGTCCGGAGAACAGTACCTCTGTGGCAC 3193
QY 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTpsArgPheLe 450
Db 3194 TCAAGGTGTCCAGCGCTGCGCTCAGGAGGACCTCAGGCTTCTGAGATTCCT 3253
QY 450 uGlnTpsAsnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTpsPheAr 470
Db 3254 GAGGTGAGTCAATGCTCCCGCTCCCAAGGAGCTGTATAGGGGGCTCTGTCTGCC 3313
QY 470 gHsTyraAspPheArgGlnSerCysTpsValArgAlaIleSerGlnGlnAspGlnLeuAl 490
Db 3314 CATTATGATTTTGAACGAGCTGCTGGGTCGATCCATATCCAGAGAGACAGCTGCG 3373
QY 490 aProCysTpsGlnAlaGlnIleSproAlaGlu 500
Db 3374 CACTGCTGGCAGCTGAACCTCGGAGAG 3404

RESULT 11
US-10-098-841-51
/ Sequence 51, Application US/10098841
/ Publication No. US20020197679A1
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Xu, Chongjun
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yundong
/ APPLICANT: Wang, Qian-Rui
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Zhang, Jie
/ APPLICANT: Qian, Xiaohong B.
/ APPLICANT: Dmanac, Radoje T.
/ TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
/ FILE REFERENCE: 784CIP2
/ CURRENT APPLICATION NUMBER: US/10/098, 841
/ PRIOR FILING DATE: 2002-03-13
/ PRIOR APPLICATION NUMBER: 09/598, 042
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/552, 317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488, 725
/ NUMBER OF SEQ ID NOS: 331
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 51

/ LENGTH: 8408
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (2859)..(6116)
US-10-098-841-51
Alignment Scores:
Pred. No.: 1, 91e-186
Score: 2133.50
Percent Similarity: 77.238
Best local Similarity: 71.458
Query Match: 71.528
DB: 13
Gaps: 2
US-10-071-838-2 (1-549) x US-10-098-841-51 (1-8408)
QY 1 MetAspValAlaGluValAlaGlySerTpsTpsAlaGlnGluAspIleIleMet 20
Db 1697 ATGACATGTTAGAGAAATGCAATGATTTGACAGGACAGAGCGGAAGACATCTATG 1756
QY 21 LysTyrgIyGlyGlyYHsArgAlaGlyLeuProGluAspIyGlyProIyProPheArg 40
Db 1757 AAGTATGACAAAGGACACAGAGCTGGCTCCAGAGGACAAAGGAGCTGAGCTGCTG 1813
QY 41 SerTyraAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProIeu 60
Db 1814 GGAATCAACAGAGCATGATGCTTTTGGCATTTTGCATAGACAGGACCTGCTCTGTG 1873
QY 61 ThrAlaArgGlnAlaLysGlnIleArgArgIleSerArgIySerIySerIyTpsValAsp 80
Db 1874 ACTGCAGGAGGCGAAGAAATTCGCGGAGATGACACAGGACAGAGTGAAG 1933
QY 81 MetLeuGlyAspTpsGluLysTyrgIyLysSerSerArgIyLeuIleAspArgAlaTyrgIy 100
Db 1934 ATCTGGAGAAATGGAGACATATTAAGACAGTACCAACATCATATGATGATGATCAAG 1993
QY 101 GlyMetProMetAsnIleArgIyProMetTpsSerValLeuLeuAsnIleGluIleMet 120
Db 1994 GGAATTCATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2053
QY 121 LysLeuLysAsnProGlyArgTyrgIleMetIySerGluLysArgIySerSerGlu 140
Db 2054 AAGTTGAAAAACCCCGGAGATTAACAATGAAAGAGGAGGCAAGAGTCACTGAA 2113
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgIyHisIlePhePheArg 160
Db 2114 CACATTCACCAATGACCTGACCTGAGAGGACGACCTCCGGAACCATGCTCTTTAGG 2173
QY 161 AspArgTyrgIyThrLysGlnArgIyLeuLeuHisIleLeuLeuAlaTyrgIyGlu 180
Db 2174 GATGATATGAGGACCAAGAGGAACTTTTATCATCTCTCTGCTTATCGAGTAT 2233
QY 181 AsnProGluValGlyTyrgIyArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyrg 200
Db 2234 AACCCGAGGTGGAGCTACAGGAGCTGAGACCAATCAACGCTTCTCTCTTAT 2293
QY 201 LeuProGluLysAspAlaPheTpsAlaLeuValGlnLeuLeuAlaSerGlnArgIySer 220
Db 2294 CTGCTGAGAGGAGGACCATTTGGCCTGCTGACGCTGCTGCGCAAGAGGACCTCC 2353
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu 240
Db 2354 CTGCGAGATTCACAGCCCAATGTTGGGACATGCGAGGGGCTCCAAAGCCAAACAGAG 2413
QY 241 HisValValAlaThrSerGlnProIySerThrMetGlyHisGlnAspIySlyAspLeuCy 260
Db 2414 CATGTGTACCAAGTCAACCAAGACATGTGCGATCAGAGCAAGAGGCTCTATGC 2473
QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGly 280
Db 2474 GGGCAGTGTGCTGTTAGGCTGCTCTCCGGAACCTGATGACGAGATCTCTCGGG 2533

```

QY 281 LeuThrLeuArgLeuThrAspValTyrLeuValGluGluAlaLeuMetProIle 300
DB 2534 CTCACCTCCCTGCTGGAGCCTGTATTTGGTGGAGAGAGAACAGGTGTGATGCCAATA 2593
QY 301 ThrArgIleAlaPheIleValGlnGlnIleValArgLeuThrLysThrSerArgCysGlyPro 320
DB 2594 ACCAGCATTTGCTTAAAGTTTCAAGAGAGAGCGCTCATGAAGCATCCAGGAGTGGCTTG 2653
QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
DB 2654 TGGGACAGCTGCTGGAGAACCAATTCCTGATACCTGGGCGCATGAACATGACACCGGTGCTC 2713
QY 341 LysHisIleuArgAlaSerMetLysLysLeuThrArgLysGlnGlyAspLeuProProPro 360
DB 2714 AAGCATTTTGGAGCTCTTAAGAAAGAACTTAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAG 2773
QY 360 ----- 360
DB 2774 GGGCCAGAGCCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2833
QY 360 ----- 360
DB 2834 GGGGAGAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2893
QY 360 ----- 360
DB 2894 ACCCTCTCTCTGGAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2953
QY 361 ----- 370
DB 2954 ACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3013
QY 370 TArgProValProAlaSerArgGlyGlyLysThrLeuCysGlyAspArgGlnAlaLys 390
DB 3014 CAGGCTGCTGGCGGCTTCAAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3073
QY 390 GProGlyProProAlaArgPheProArgProIleTrpSerAlaSerProProArgAlaLys 410
DB 3074 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3133
QY 410 GArgSerSerThrProCysProGlyGlyValAlaValArgGluAspThrTrpProValGlyTh 430
DB 3114 TCGTTTTCACAGCCCTGCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3193
QY 430 TGlGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnGlySerTrpArgPheLe 450
DB 3194 TCAGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3253
QY 450 GlnTrpAsnSerMetProArgLeuProThrAspLeuAspValGluGlyProTrpPheArg 470
DB 3254 GAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3313
QY 470 GHisTrpAspPheArgGlnSerCysTrpValArgAlaIleSerGlnGluAspGlnLeuAl 490
DB 3314 CCATTATGATTTTGAACGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3373
QY 490 AProCysTrpGlnAlaGlnHisProAlaGlu 500
DB 3374 CACCTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3404

```

```

RESULT 12
US-10-098-841-50
Sequence 50, Application US/10098841
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Jian-Rui

```

```

APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhimei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20020197679A1e1 Nucleic Acids and
FILE REFERENCE: 784CIP2
CURRENT FILING DATE: 2002-03-13
PRIOR FILING DATE: 2000-06-20
PRIOR FILING DATE: 2000-06-20
PRIOR FILING DATE: 2000-06-20
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 331
SOFTWARE: pf_FL_genes Version 1.0
SEQ ID NO 50
LENGTH: 8420
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2859)..(6128)
US-10-098-841-50
Alignment Scores:
Pred. No.: 1,916-186 Length: 8420
Score: 2133.50 Matches: 408
Percent Similarity: 77.23% Conservative: 33
Best Local Similarity: 71.45% Mismatches: 58
Query Match: 71.52% Indels: 72
DB: 13 Gaps: 2
US-10-071-838-2 (1-549) x US-10-098-841-50 (1-8420)
QY 1 MetAspValAlaGluValAlaGlySerTrpAlaGlnGluArgLysAspIleLeuMet 20
DB 1697 ATGACATGATGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1756
QY 21 LysTrpGlnLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
DB 1757 AAGTATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1813
QY 41 SerTyrAsnAsnAsnValAspHisIleuGlyIleValHisGluThrGluLeuProProLeu 60
DB 1814 GGAATCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1873
QY 61 ThrAlaArgGluAlaLysGlnIleArgArgGluIleSerArgLysSerLysTrpValAsp 80
DB 1874 ACTGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1933
QY 81 MetLeuGlyAspTrpGlnLysTrpLysSerSerArgLysLeuIleAspArgAlaTyrLys 100
DB 1934 ATGCTGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1993
QY 101 GlyMetProMetAsnIleArgGlyProMetTrpSerValLeuLeuAsnIleGluGluMet 120
DB 1994 GGAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2053
QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGlnGlyLysArgSerSerGlu 140
DB 2054 AAGTTGAAGAAACCCCGAGAGATACCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 2113
QY 141 HisIleGlnArgGlyLeuAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 2114 CACATCCACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2173

```

```

Qy 161 ASPARGLYRGLYTHRLYGLINARGULLEULEUHSILEULEUALATYRGLINUTYR 180
Db 2174 GATCGATATGAGCCAGACAGAGGAGACTATTTCTACATCTCTGCGCTATTCGAGATAT 2233
Qy 181 AENPROGILUVALIGLYTYCYASARGAPLEUSERHSILEALALEPHELEULEUTYR 200
Db 2234 AACCCGAGGTGGGCTACTGTCAGAGACCTGAGCCATCCACCGCTTGTCTCTCTTAT 2293
Qy 201 LEUPROGILUGLULASPRALAPHETPALALEUVALIGLEULEUALASERGLUATGHSER 220
Db 2294 CTGCTGAGGAGAGACCCATTTCTGGGCTACTGGTGCAGCTGCTGGCCAGTGAAGGACTCC 2353
Qy 221 LEUGLNGLYPHEHSERPROASNGLYGLYTHRALIGLNGLYLEUGLNASPGLNGLNG 240
Db 2354 CTGCGAGATTCACAGCCCAATGGTGGGACGTCAGGGGCTCCAGAACCAAGAGAG 2413
Qy 241 HISSVALVALALATHSERGLNPROLYTHMETGLYHISGLNASPLYLEUASPLEUCYS 260
Db 2414 CATGGGTGTAACCAAGTCCAAACCCAGACCATGTGTCATCAGGACAGAAAGTCTATG 2473
Qy 261 GLYGLNYSERPROLEUGLYCYGLLEULLEATGLILEULLEAPGLYILESERLEUGLY 280
Db 2474 GGGCAGTGTGCTGCTGTTAGGCTGCTCTCCCGAACCTGATTGACGGATCTCTCGGG 2533
Qy 281 LEUTHLEUARGLEUTTPASPRVALTYRLEUVALIGLUGLUGLINALALEUWETPROILE 300
Db 2534 CTACCCCTGCGCTGTGGGACGCTGATTTGGTGAAGAGAAACAGGTGTGATGCCAATA 2593
Qy 301 THRARGILEALAPHELYEVALIGLNGLYLEUARGLEUTHRYLTHSERARGYNGLYPRO 320
Db 2594 ACCAGCATTTGCTTTAAGTTCCAGCAAGGCGCTCATAGAACATCCAGCGTGTGCGCTG 2653
Qy 321 TTPALAAARGHECYASARGPHEVALASPTHTTPALAAARGPULASPTHTVALLEU 340
Db 2654 TGGGACGCTGCGGACCAATTTCTGATCTCTGGGCTGAGTGAAGACACCGTGTCTC 2713
Qy 341 LYSHSILEUALASERMETLYLYLEUTHRYARGLYSGLNGLYASPLEUPROPROPO 360
Db 2714 AAGCATCTTAGGCTCTCTAGAAACTTAACAGAAAGAGGAGGACCTGCCACCCCA 2773
Qy 360 ----- 360
Db 2774 GGCCCAACAGCCCTGGAGAGAGGTGTGGCAGAGAACCCCGACGATCGAACCTTG 2833
Qy 360 ----- 360
Db 2834 GGGGAGTCCAGAGCCACCCACCATGCCCACGCGTTCCCATGCCAGGACACAC 2893
Qy 360 ----- 360
Db 2894 ACCCTCTCTGGGATCAGAGACTAAGGCGTGTCTGATGACACCAAGGGGCTC 2953
Qy 361 ----- 361
Db 2954 ACACAGAGACCCCAAGACTCCAGAGATGCGAAGACGAGAGAGGGCTCTGGGACAC 3013
Qy 370 RARGPROVALPROALASERARGLYGLYLYETHRLEUCYELYSGLYASARGHIALAPR 390
Db 3014 CAGGCTGTGGCGGCTTCACTGGTGGGAGAGACCTCTGCAAGGGATATAGGACAGCC 3073
Qy 390 OPROGLYPROPROALAAARGPHEPROARGPROILETPSERALASERPROPROARGLA 410
Db 3074 TCCAGGCGCACAGCCGATTCAGCGGCCCATTTGCTCAAGCTTCCCGCATGGGATC 3133
Qy 410 OARGSERHTHRPROCYSPROGLYGLYALVALALARGLULASPTHTYRPROVALIGLY 430
Db 3134 TCGTTTTTCCAGCGCTGCTGTTGGGCTGTCCGGAGACACGTAACCTGTGGGACAC 3193
Qy 430 RGLNGLYVALPROSERPROALALEUVALIGLNGLYPROGLNGLYSERTPHARGPHELE 450
Db 3194 TCAAGGCTGTGCCAGCGCTGCTGCTCAAGGAGACCTCAAGGATCTCGAGATTCCT 3253
Qy 450 UGINTPANSERMEPROARGLEUPROTHRASPHEUASPRVALIGLNGLYPROTPHARG 470

```

```

Db 3254 GAGGTGAGTCAATGCCCCCGGCTCCCAAGGACCTGGATATAGGGGCGCTGGTCCC 3313
Qy 470 GHSITYPASPRPHEARGGLNLSERYSTTPVALARGALAZISERGLNGLUASPRGLNLEU 490
Db 3314 CATATATATATTTTAAACGAGCTGTGGGTCTGCTGCATATCTCCAGAGGACGACTGCC 3373
Qy 490 APROCYHTPGLNALAGLUNHSIPROALAGLU 500
Db 3374 CACTGCTGGCAGCTGAACAACCTGGGAGAG 3404

RESULT 13
US-09-764-869-1824
; Sequence 1824, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1824
; LENGTH: 9805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1824

Alignment Scores:
Pred. No.: 1,58e-128 Length: 9805
Score: 1507.50 Matches: 487
Percent Similarity: 20.97% Conservative: 6
Best Local Similarity: 20.71% Mismatches: 17
Query Match: 50.54% Indels: 1844
Gaps: 11

US-10-071-838-2 (1-549) x US-09-764-869-1824 (1-9805)
Qy 24 LYSGLYHSIARGALAGLYLEUPROGLULASPRLYSGLY-----PROLYSPROPHARGSER 41
Db 2492 CAGGGCTCCAGAGCCCAAGACCCAGACATCCGCGGCGGCTCTGGAGGCTCGAGCTCC 2551
Qy 42 TYRASNASPRVALASPRHSILEUGLYLILEVALHSILEUTHRLLEUPROPROLEUTHR 61
Db 2552 GCTAATCTCAACATGCTCATTTG----- 2575
Qy 62 ALAARGLUALALYS-GINTLEARGTGLULISERARGLYSERLYSTTPVALAPME 81
Db 2576 -----ACAGCAAAATTCGCGGAGATCAAGCCGAAAGACAAAGTGGGTGATAT 2623
Qy 81 ELEUGLYASPTTGLULYLYTYRLYSERSEARGLYS----- 93
Db 2624 GCTGGAGACTGGGAGAAATAACAAAAGCAGAAAGGTAACTGTGAGAGAGAGAAACA 2683
Qy 93 ----- 93
Db 2684 CTCTCTGCAAGACAGGGGACAGGACCCATGCTGTGTGCTGGCAGCATGAGGCTCTC 2743
Qy 93 ----- 93
Db 2744 AGAGGGTGGCGGACACTGCTGCTGCCAGAGACTGAGGCTGTGCGCAATTTCC 2803
Qy 93 ----- 93
Db 2804 TGCCTATTGCGAGAGCTCACTTGCAGAGGAGGAATCTGAATCTAGGCTGGACTAC 2863
Qy 93 ----- 93
Db 2864 CCGAGCTCAAGGCTAGGAGATGCTGTGTGACTGAAGAGAAAGGTTGAGATCAGA 2923
Qy 93 ----- 93

```

Db 2924 GTTTCAGTCTGAGTGTCCATCCACTCTTCACTCTGGGAAAGAGACCTGTCTCCAGC 2983
QY 93 ----- 93
Db 2984 TTGATCTACCTCTACTGAGAAATCATGGGGCCAAAACCAATTTCCAGAAATCCCGG 3043
QY 93 ----- 93
Db 3044 GCTCTGCTCTCACTGGGGTCAACCCCGTGGCTGTGACACAGATTTTCTGCCACA 3103
QY 94 -Leu1leapArgal aTyLysGlyMeProMetAsn1leArgGlyProMetTrpSerVa 113
Db 3104 GCTCATAGATCGAGCATACAGAGGAATGCCATGAAATCCGGGGCCCATGTGTGAGT 3163
QY 113 -Leu1leuAn11eGluGluMeLysLeuLysAsnProGlyArgTyTrpGln1le----- 130
Db 3164 CCTCTGAAATGAGAAATGAAATGAAAAAACCCTGGAGATACAGGT -ACGCTCAG 3222
QY 130 ----- 130
Db 3223 CCAGAGCACAAACAGAGACAGCCGTGTGGGGCCCAAGTCTCCAGCTGAGAGGAAAG 3282
QY 130 ----- 130
Db 3283 TCAAGACACCTGGGGAGCTGGGGGTGAAGGTCAATGAAACCTGGGGCAAGATGTT 3342
QY 130 ----- 130
Db 3343 GACACAGTCACACAGACAAACTCAGCTGTGTGACCTCTGGCTTGATTAACAAGCC 3402
QY 130 ----- 130
Db 3403 AAAATGACCTTTCTGCAAGAAAGAACTTCTTCTGTCTCTTCCGAAAGTCTGAC 3462
QY 130 ----- 130
Db 3463 TGTGGGCTGACTGCCACTGGGGGAGAGAGTCTTCATCTGTTCTGAGCTGCTTCTCC 3522
QY 131 -----MetLysGluLysGlyLysArgSerSerGluHis11leG1 143
Db 3523 GCTTGGCCCTGCTCCCTACAGATCATGAAAGAGAAAGGCAAGGTCTCTGAGGACATCCA 3582
QY 143 nArG11eAPArGaPeVa1SerG1YThreLysArgLysHis11lePhePheArgAspArgTy 163
Db 3583 GCGCATCGACCGGACATAGCGGGACATTAGAGAAACATATGTTCTTCAGGGATGATA 3642
QY 163 rG1YThrLys----- 166
Db 3643 CGGAACCAAGTAAGCTTACGGGAGCCACAGGATCCACAGAGATGGGGTGAATGAGAG 3702
QY 166 ----- 166
Db 3703 GATGGGGGCTTCCCGGAGCAGAAAGCCAGGGTCAACCCAGAGGATGACAGCTGCCAA 3762
QY 166 ----- 166
Db 3763 GAGCTCTCCCGGCCAGGAGACCGCGACCATGAACGAGACCTCCCTGTTCGAAG 3822
QY 166 ----- 166
Db 3823 CCTGGGCGAAGTGSAAATGTGGGGCCAGAACCCAGAGAGATCTCTGAGAGATGGAAG 3882
QY 166 ----- 166
Db 3883 GCAGCAACAAATCATGCAATGTGAAGGTGCTCTCTGACCATGGGAGCCCAT 3942
QY 166 ----- 166
Db 3943 GGTAAGACCCAGGAGGGGTGCGAGATGAGGGGCCATGAGCCCCCAGGCAACAGTG 4002
QY 166 ----- 166

Db 4003 ACAGCACCAATGTGGAGAAATTAGGGGTCTTGAAACTCTATCCAGGTCCGTGGGA 4062
QY 166 ----- 166
Db 4063 ACATGACATGGACAGCCACGTTGGCAGCAGTGGGAGTGGCTTCACAAAGCTGATGG 4122
QY 166 ----- 166
Db 4123 ACTTGAACACATCCCAAGTGTCAAGATTTGAACCACTGATTTGGAACTGAC 4182
QY 166 ----- 166
Db 4183 ATCCATGAAACACAGATGCCAGGTTCACCTGTAATCTCTGACTCACACAGGAG 4242
QY 166 ----- 166
Db 4243 CCTTGGGGAGCGCTTCAACAGGGAATGGGAGAGCAAGGTGTCTCTCTTCAAC 4302
QY 166 ----- 166
Db 4303 GGAAGACCCAGTGAAGAAAGGAGACAGCCGTGTATGCCGACAGACGTGGGTGATCC 4362
QY 166 ----- 166
Db 4363 TAGATGATTTTGTGAGGAGACAGAACCCAAATTAAGTACACAGATGATGCC 4422
QY 166 ----- 166
Db 4423 ATTCCTAGGCATTTCTGAAAAAGCCAAACACAGGAACTGAAAGCATCTGGGTGCC 4482
QY 166 ----- 166
Db 4483 AGGGGCTGACGATGGGGAGAGGCTGGGTGATAGGGGCCACCTGAGACTTGGAGGA 4542
QY 166 ----- 166
Db 4543 TGAAGAGTCCGCCAGAGGGGCTGAGCGGTGCCGGAGACTTGACATTTGTTTG 4602
QY 166 ----- 166
Db 4603 GAACCGTGAGAACTGTACACACACAGACTGAACCTGGCTGTGCAAACTGAAAAAA 4662
QY 166 ----- 166
Db 4663 AAAAAAATCATTCAGAGTGAAGAAAGATCAGGCAAGTCACTGACAATGGCTATTGC 4722
QY 166 ----- 166
Db 4723 ATGTCAAGATGTGATTTTACTGAACAATTTCTTCAAGAGTCTGAGCCCTGAAGAGCT 4782
QY 166 ----- 166
Db 4783 CACTGCTTATCTGTGAAACATCTGAACCTGAATGGATTTGCTGTAGCTTTGTAGA 4842
QY 166 ----- 166
Db 4843 CAAAGTGAATTAACAACATCTGCACAAAACAAACCAAGCCCCCTTCTGTGTTCTTA 4902
QY 167 -GlnArgGluLeuLeuHis11leLeuLeu1aTyTrpGluLysPro----- 182
Db 4903 GGCAGCGGAACTACTCCACATCTCTCTGGCATATGAGAGATTAACCCGGTGAATTC 4962
QY 182 ----- 182
Db 4963 CCGGAGTGAAGTCCCGGGCCATATTTCCATATTTGACAGAGTGGGTCTGTGGGGG 5022
QY 182 ----- 182
Db 5023 TGTGTTGCTTCTTTAAAGTTAGTATTTGACCAACAGGATTAAGAGGTAGAGATGT 5082
QY 182 ----- 182
Db 5083 CAGTCAACCGCTGCAATAAATCTCAAGAAAGGGGTGTCTCAAGGGGTCAAGCTGAGA 5142

```
OY 182 ----- 182
Db 5143 CACAAAGAGTACGAGGCCCGAGCTCTGGTGTCACTGGGCTGACCAACACATTTTCAGA 5202
OY 182 ----- 182
Db 5203 ACAAGAAATGACGCCCTCTCTGGGGGTGCCCCAAGGCCAGAGCTTGGCAGCATCCG 5262
OY 182 ----- 182
Db 5263 ACACAGATGTGCTATACAGACATTTTGAACAAGTGTGAAGTGCCTGATGACTT 5322
OY 182 ----- 182
Db 5323 GGCCTTGTATGAAATGATGTGCATCTGAGGAAGCCTTTTTCAGAGGAAGCCTCT 5382
OY 182 ----- 182
Db 5383 CTTTCAGAGGAAGCCTCTCTCACTGACCTTGCCTCTCCATGACATGAGTCTCCAGG 5442
OY 183 ----- 183
Db 5443 TGACTCAGGCGCTCCAGGAGTATGCTTCATGTGACTGTGGCTTTGAGAGAGGTGG 5502
OY 185 ----- 185
Db 185 1YTCYsArpAspLeuSerHis11e1a1a1aLeuPheLeuTyrLeuProGluGlu 205
5503 GCTACTGCAAGGACCTGAGCCACATCGCCGCTGTGCTCTCTATCTCTTGAGAGG 5562
OY 205 sPAlaPheTPrAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGln----- 222
Db 5563 ATGCATTCTGGGCACTGGTGGAGCTGCTGGCAATGAGGCACTCTCTGCA-GGGTAAg 5621
OY 222 ----- 222
Db 5622 TGAACAGCTGCCCGGGAGCCTCTCTGACGCCAGACTGGGATGGCCACCTGGCGGGT 5681
OY 222 ----- 222
Db 5682 GATCAGACTTTTCAGCCAAAGGCAACCTCTTGTGTGCCAGCTGTTGGAGACTTAAg 5741
OY 222 ----- 222
Db 5742 ATGCTCTGTCTGAGGGTCCCAAGAGTCCACGGCTGACCCCAAGGCCCAATGAGAG 5801
OY 222 ----- 222
Db 5802 CTTCTATCCCAATCAGACAGAGGGCATCTCATCTCCCGTGGCCACCTCTGTGTCTG 5861
OY 222 ----- 222
Db 5862 GAGCAGCGCCCTCGGCTCTGATTTCTGTGAGCTGACTCTCCCTCTGAGAGTCTCC 5921
OY 222 ----- 222
Db 5922 TGCCCTCAGCTGCCGGGCTCTGTGTCATCGGTGCCACGAATGGCCGACCAAGCC 5981
OY 222 ----- 222
Db 5982 CAGGTGGAGCATCTTCCCAATCCCTGTCTGTCCCTGGCCGACCCCACTACAGAGATGA 6041
OY 222 ----- 222
Db 6042 CCGGAGAGCCAGCGCCCAACCAATTCGAGCGCGCTGTGTGGCTTGAAGTCAAGGCTT 6101
OY 222 ----- 222
Db 6102 GCCCTTTTTCACCTGGCCAGAGGCTTCCAGGGGAACCTCAGCAGAGGCTCCAGGGA 6161
OY 222 ----- 222
Db 6162 ATGTTCCGCCCCCACTTCCCAAGGTTAAGCCGCATGTTGGGGTCAACAGATGGAGGG 6221
OY 222 ----- 222

OY 222 ----- 222
Db 6222 TGGAGTAGCCTTGGGGTTTGGGGGCTCTCCAGCTGCCAGCTTTCAGCTGATGCT 6281
OY 223 ----- 223
Db 6282 CCACATCTTGGGGGAAGCTGTGATTTCAATGATGGGCTGGGGGCTTTCAGAGATTTCA 6341
OY 226 ePProAsnGlyThrValGlnGlyLeuGlnAspGlnGlnGluHisValValAlaThrs 246
6342 GCCCAATGGCCGAGACCGTCCAGGGGCTCCAGACCAAGAGCATGTGTAGCCACGT 6401
OY 246 eRGlNProLysPheMetGly----- 252
Db 6402 CACAAATCCAAAGCAATGGGGCATGAGTGAATTATGTCTCCCTCAGCTCTTCCAGAG 6461
OY 252 ----- 252
Db 6462 CCTGCTCCGCTGGGGCTGTAGAGAGAGGGGGCTGGGGCCCTGTGTGGGCTGTGAC 6521
OY 252 ----- 252
Db 6522 TGGCTGAGTCCAGCCAGGGCTGACCTGGAGAGTGGGTTTCATGGGCTGGAGTTG 6581
OY 252 ----- 252
Db 6582 GTTTCCTTCTGCTGCTGAGAGAGACAGAGGACAGGGATGGGGCCAGCTCCGACAG 6641
OY 252 ----- 252
Db 6642 GCAGGGCAAGGCAGTGTGTCCACCGGAGATGTGGGAAGTGAAGTGTGTGGGAGC 6701
OY 252 ----- 252
Db 6702 TCTGACACCGCCAGTGTCTGTCACTAGGGGAAGGTTTCAGAGACCCGAGAAGGG 6761
OY 253 ----- 253
Db 6762 AGGTTTATGGGAGCGCAGGTGCTGAGCACCTGTGTGCTTCATGAGCAAGAAAG 6821
OY 258 sPLeuCyGslYgInCySerProLeuGlyCyBlau11eArg11aLeu11eAsp-Gly--- 276
6822 ATCTATGTGGGAGGTGTTCCCGTTAGGCTCTCATCTCGGAATGATTAAGCGGGTAA 6881
OY 276 ----- 276
Db 6882 GAGGATAGGAGACCTGTGCTCAGGAGCCTTCTGTGCTGTGAGTGGCTTCTCC 6941
OY 276 ----- 276
Db 6942 AGCCGGGGGTCTGGCTCATCTCCAGCCACAGAGGCTCAGCGGGGTCCCAAGAGACA 7001
OY 276 ----- 276
Db 7002 CACAAAGAAAACCTTGTCCCAAGAGGGGTCAATCCAGGGCAATGGCTGGGGCTCAGGCC 7061
OY 276 ----- 276
Db 7062 CAGCTCATGGGAGACATGGGCCAGACCCGACTTGAGAGGCTCAGGAGACCTCAAGC 7121
OY 276 ----- 276
Db 7122 CTGGGAGAGCCCTCTTCTCCAGAGCCACATCTCCCACTCAATATGATGCCCCCATGAG 7181
OY 276 ----- 276
Db 7182 GAGCTTCAAGACTTGTCTGACCCAGCGTCTTGAGGGCTCAGGCGACCTCATGGGGAA 7241
OY 276 ----- 276
Db 7242 GGTCACTGACTTGGAGCTGAAGCCCAAGTGTGGCAGCTCGAGCAACAGGCCCGCAGCC 7301
OY 277 ----- 277
Db 7301 ----- 277
OY ----- 283
Db ----- 283
```



```

Db      7302 TGGAAAGACAGATTCTTTACACCTGCTGTCCCAAGATCTCTCTGGGCTCACCCCTG 7361
Qy      284 ArgLeuTTPAspValTylLeuValIgluGIgluGlnAlaLeuMetProIleThrArgIle 303
Db      7362 GCGCTGTGGAGAGCTGTATCTGTGTAGAGCGAAGCGGCTGTATGCGCATTAACAAGAAATC 7421
Qy      304 AlaPheIValIgluGln----- 309
Db      7422 GCTTTAAGTTACAGACAGTAAGTCTACGTGTGCCAGCGGGGCTGGGGAGCCCTGGG 7481
Qy      309 ----- 309
Db      7482 GTCAAGACCCCGAGCTGGCCGAGGGGAGCTTCTCTCAACATGTCTCATATTCGGTGTCT 7541
Qy      309 ----- 309
Db      7542 GGCACAGAGGAGGTCCGGCCAGGTGGCTGGGCAAGACATGTGACACCGAGCCCATCC 7601
Qy      309 ----- 309
Db      7602 CTCACATGATCCGACAGGAGAAAGTGTCTACCACTCTGACTTTTCATCTGGGTCCCGAG 7661
Qy      309 ----- 309
Db      7662 CCACAGTCTCTGTGTATATCTGACACCTGGGGTGGCCACAAGAAAGATCCGGCACCGAC 7721
Qy      309 ----- 309
Db      7722 CAGTAGAGACTGAAGTGCCACGGGGTATGAGCTGTGACCATTCACAGTAATCCCT 7781
Qy      310 ----- 781
Db      7782 GGCCTGATATCCACCTGTCTCTTAGAGCGCTCTACAGAAAGCGTCCAGGGTGGCGCGTGG 7841
Qy      322 AlaArgPheCyAsnArgPheValAspThrTAlaArgAspGluAspThrValLeuIys 341
Db      7842 GCACGTTTTTGGCAACGGGTGCTGTGATCTGTGGCCAGGATGAGAGACCTGTGTCAAG 7901
Qy      342 HisLeuArgAlaSerMetIylValLeuThrArgIylGlnIylAspLeuProProAla 361
Db      7902 CATCTTAGGGCTCTATAGAAACTAACAGAAAGACAGGGGAGCTGCCACCCCAAG- 7960
Qy      361 ----- 361
Db      7961 TGGGCTCAGTGCATGTCCCTCTCCATGTCACTCTGGGGTACTCAGTAGAGGGAG 8020
Qy      361 ----- 361
Db      8021 TGCCCGGAGACCCGCAACCTTACTACTGGGCTTCTCTTCACTTTTCTCTCTCTT 8080
Qy      361 ----- 361
Db      8081 CCTCTGAGCTCTAAGAAAGTACAGAGGCGCACCGGCTCTCAGGGAGAGCGCTCAGTGC 8140
Qy      361 ----- 361
Db      8141 GTGTACTGACATGCTGTGCAAGCAGAGGGGAGATGTGGCAAGACCTTCAACAAAGC 8200
Qy      361 ----- 361
Db      8201 CCCCTCCACATTTCACGGGTGTCTCTCTCCCTCGAGAGGCGCTCCAAATTACTGA 8260
Qy      361 ----- 361
Db      8261 CGAGCCAGAGCCATTGTGGAGAACCCGCGCTCTCCCTGCAAGACCCACAGCCTCAGA 8320
Qy      361 ----- 361
Db      8321 GAGCAGCAGAGGCCCTCTACTCTGTGACGCTCTCCAAAGTTTGCAGAGCAAGAAAGCTTG 8380
Qy      361 ----- 361

```

```

Db      8381 GAGCCAGGAGACAAAGGAATCCGTGTCCCTGACCCACAGAGCATTTCAAGGAGAGGCC 8440
Qy      361 ----- 361
Db      8441 AGAGCCAGAGCCAGAGATTTCAGCCAGAAAGTGGAAAGGTCAGTCTCTGGCATGACTGGGC 8500
Qy      361 ----- 361
Db      8501 AGCCAGAGAGGAGAGGGTACCCACGTCCGGGCCAATACCCACTGGCGGAGACGGGT 8560
Qy      361 ----- 361
Db      8561 CCCCATGTAGAGTGAACAAGGGCTGGGTGACATCCAAAGCCCTTCCACTGAGTTCTGA 8620
Qy      361 ----- 361
Db      8621 CTGGGGGCTGTATCCAGGCCCAACAGCCCTTGGAGCGAAGGTGTGTGCGAGAGAGCCCC 8680
Qy      361 ----- 361
Db      8681 AGCCAGTGAACCTGGGGGCGATGCCAGAGAGCCACCGGCATGCACAGAGCTTCCC 8740
Qy      361 ----- 361
Db      8741 CACGCCAGGACATGACACCCCTCTGGGATCAGAGACTACAGGCGTGTCTCGGT 8800
Qy      361 ----- 361
Db      8801 GTCAGGCCAGGGGGCCACACAGGAGCCCGAGAGCTCCGAGATGCAGGAGTGGGGGCC 8860
Qy      361 ----- 361
Db      8861 CAGCCCGGAAAGCCCTGGGTGGGCTCACTGAGATGCTGACCGGCTGTCTTCTTCA 8920
Qy      362 ---LysProIgluGlnIylSerSerAlaSerArgProValProAlaSerArgIylIyls 380
Db      8921 GCCAAACCCGAGCAAGGGTGTCTGTGGCATTCAGGCGCTGTGGCTTCAAGTGGCGGAG 8980
Qy      381 ThrLeuCyIylGlnIylAspArgGlnAlaProProIgluProProAlaArgPheProArgPro 400
Db      8981 ACCCTCTCAAGGGGAGAGAGGAGGCGCTCTCAGGCCACCAAGCCGGTGTCCCGGGGCC 9040
Qy      401 IleTPSerAlaSerProProArgAlaProArgSerSerThrProCyAspProIylIylAla 420
Db      9041 ATTGTGAGCTTCCCGCCACGGGACCTGTCTTTCACACCTGTCTGTGGGCT 9100
Qy      421 ValArgIylAspThrTyrProValIgluIylGlnIylValProSerProAlaLeuAlaGln 440
Db      9101 GTCCGGAAGACACCTTACTGTGGGCACTCAGGGGTGTGCCAGCGCGGCTGTGAG 9160
Qy      441 GlyIylProIgluIylSerTTPArgPheLeuGlnTTPAsnSerMetProArgLeuProThr 460
Db      9161 GGAGGACCTCAGGGTCTCTGAGATTCCTGAGTGAATCTCATATGCCCCGCTCCCAAG 9220
Qy      461 AspLeuAspValIgluIylProTTPPheArgIylIylAspPheArgIylSerCySerTTPVal 480
Db      9221 GACCTGGAGCTGAGAGGCGCTTGGTCCGCAATTAATTAATTCAGACAGAGCTGTGGGTC 9280
Qy      481 ArgAlaIylSerGlnIylAspGlnLeuAlaProCySerTTPGlnIylGlnIylProAlaGln 500
Db      9281 CGTGCCATATCCAGAGAGACAGAGTGGCCCTGTGGCTGGAGCTGAACACCTTCGGAG 9340
Qy      501 ArgValArgSerAlaPheAlaAlaProSerThrArgPheAspArgIylIylThrProPheArg 520
Db      9341 CGGGTGAATGCGCTTGTGTGACCCAGCACTGATTCGACAGGAGACCCCTTCAGA 9400
Qy      521 AlaArgAspGlnIgluIylProCyAlaProThrSerGlnIylProCyIylLeuCyIylLeuIylSer 540
Db      9401 GCTAGGAGCAAGACAGCATGTGTCTCCACTCAGGCGCTTGTGTGGGCTTCCACTTG 9460
Qy      541 GluSerSerGlnPheProProIgluIylPhe 549
Db      9461 GAAAGTTCTCAGTTCTCTCCAGGCTTC 9487

```

Alignment Scores:
Pred. No. .

Score:	1.58e-128	Length:	9805
Percent Similarity:	1507.50	Matches:	487
Best Local Similarity:	20.97%	Conservative:	6
Query Match:	20.71%	Mismatches:	17
DB:	50.54%	Indels:	1844
	10	Gap:	

24 [vsg]w: - 1-7857 (1-9805)

24	LySGIYHSATGVALAQLYENPProGIunAplysGIY-----	ProLYsPProPheArgser	41
2492	CAGGGCTCCAGAGCCCAAGACCAGATCCGCGGGCGGCTTCGGAAAGCTCCAGACTCC		2555
42	TyASAsAsAsnValAspHisLeuGIYLYIleValHisGIunthrGIunleuProProleuThr		61
2552	GCTAACTCAACACTGCCTCATTTG-----		2575
62	AlAaArgGIunAlaLYs-GInIleAArgArgGInIleSerArgLYsSerLYsTPPValAspMe		81
2576	-----ACAGCAAAATTCGGCGGGAGATCAACCCGAAAGACAAAGTGGTGATAT		2623
81	ILeuGIYAsPTTPGInLYsTYrLYsSerSerArgLYs		2623
2624	GCTGGAGACTGGAGAAATTCAAAGACAGAAAGTAACGTGTGAGAGAGAAAGCA		2683
93	-----		2683
2684	CTCTCTGCAGACAGAGGACAGGCCACCCATAGCTTGTGGCTGTGCACCATCAGCCTCTC		2743
93	-----		2743
2744	AGAGGGTGGCGGCACTGTCCTCTGCCAGAGGACTGCAGGCTGTGTCAGATTCC		2803
93	-----		2803
2804	TGCCTATTCTGTGCAGCGTCACTTTGAGAGGAGGAATCTGAATCTAGGGCTGGGAATAC		2863
93	-----		2863
2864	CCGAGACTCAAGCTAGAGATGCCCTGTGACTGAAAGAAAGAAAGTTCAAGTACAGA		2923
93	-----		2923
2924	GTTTCGACTGTGAGTCCATCTCACTCTTTCAGTCTGTGGAGAGAGACCCTGTCCAGC		2983
93	-----		2983
2984	TTGATCTCACTCTACTAGAGAAATCAGGGCCAAAAACAACAATTTCCAGAAATCCCGG		3043
93	-----		3043
3044	GCCTGTGCTCTACTGGGGTCAACCCCGCTGTGACACAGATTGTTTCTGCCACA		3103
93	-----		3103

Dy		94	-Leu11leAspaRga1eTyrlYsglYmetProme1asn1leaRgylYPrometTrpSerVa	113
Dd		3104	GCTCAATAGATCGACATTCAGAGGAATCCCATATAACATCCGGGCCCCGATGTGGTCACT	113
Dy		1113	lleuleAuan1legluglUmelYsleuYsaSPProglYARgTYrglnle-----	3167
Dd		3164	CCTCTGAACATTGAGAAATGAAGTTGAAAAACCOCGAAAGTACCAAGT-ACGCTCAG	3222
Dy		130	-----	130
Dd		3223	CCAGAGCAAACAACAGAGCAGGCCGTGTCCGGGCCAGGTCCTCACGTGAGGAACG	3282
Dy		130	-----	130
Dd		3283	TCAAGACCACCTGGGGAGCTGGGGGTGAAGSTCAGATGAACAACCTGGGCACAGATGT	3342
Dy		130	-----	130
Dd		3343	GACACAGTCAACAACAACACTCAGCTCTGTGACCCCTCGCTTAGTAACAAGCC	3402
Dy		130	-----	130
Dd		3403	AAAATCAGCTTTCGAGAAGAAACCTTCCTTGTCCTTCCTCCGAAAGTCTGAC	3462
Dy		130	-----	130
Dd		3463	TGTGGCTGATCGCCACTGGGGCAGGGAGTCTTCATCTGTTGAACTGCTTCTCC	3522
Dy		131	-----	131
Dd		3523	GCTTGGCCCTGCCCTTACAGATCATGMAAGGAGGACAAAGGTCATCTGACACATCCA	3582
Dy		143	NARG11easPaRspValSerGlYThrleuAryLsh1sllePheheaRgspargTy	163
Dd		3583	GGCATTGACCCGGACATZMACGGGACATTNAGAAAGCATATGTTCTTCAGGATCGATA	3642
Dy		163	RGLYThrlYs	166
Dd		3643	CGGAACCAAGTAAAGCTTACGGGACCAACAGGTCCTCAGCAGAGTGGGTGATAGAGG	3702
Dy		166	-----	166
Dd		3703	GATGGGGCTTCCCAGAGACAAGAACGAGGTCACCCAGAGAGATGACACAGCTGCCAA	3762
Dy		166	-----	166
Dd		3763	GAGCTTCCGGGCCAGGAGACGCCGGACCATTAACGAGCACTTCCTGGTTCMAAG	3822
Dy		166	-----	166
Dd		3823	CCCTGGGCCAGCTGSAACATGTGGGGCCAGAACCCAGAGAGATCTTGAGAGATGAAG	3882
Dy		166	-----	166
Dd		3883	GCAGAAACAATAATANTGCACATGTGAAGGGTGTCTTCCTGACCCATGGGACCCAT	3942
Dy		166	-----	166
Dd		3943	GGTAGAACCCACGGAGAGGTGGCAGATTAAGGGCCCATGAGCCCCCAGCAACAGTG	4002
Dy		166	-----	166
Dd		4003	AACAGACCAAAATCTGTGGAGAAATTAGGGGTCTTGAAACTCTCATCCAGGTCCGCTGGGA	4062
Dy		166	-----	166
Dd		4063	ACATGACATGGCAGACCACTGTGGCAGCCAGTTGGGCAGTGCTGACAAAGCTGATGG	4122
Dy		166	-----	166
Dd		4123	ACTTGAACCAACATCCCAAGTGTCAAGATATTGAACCACTGATTTGGAAACTGAC	4182

QY 166 ----- 166
Db 4183 ATCCATGAAACAGCATGCCAGTTCACTGCTTGACTCTCTGCTACACACGAG 4242
QY 166 ----- 166
Db 4243 CCTGGGGGACGGCTTCAACAGGGATGGGGAGACAAAGCTGTCTCTCTCAAC 4302
QY 166 ----- 166
Db 4303 GGAGAGCCAGTGAAGAAAAGGAGCGCGGTGATGCCGACGAACTGGGTGATCC 4362
QY 166 ----- 166
Db 4363 TAGATGCAATTTGCTGAGGAGACAGACCCTCAATAGCTACACAGTAGATTCCC 4422
QY 166 ----- 166
Db 4423 ATTCTAGGCAATCTGAGAAAAGCCAAACACAGGAGCTGAGAGCACTGGGTGCC 4482
QY 166 ----- 166
Db 4483 AGGGCTGACGATCGGGAGAGGCTGGGTGATAGGGCCACCTGAGACTTGAGAGA 4542
QY 166 ----- 166
Db 4543 TGAAGAGTCGCCCCAGAGGGGCTGAGCGGTGGCCGGAGACTTGACATTTGTTG 4602
QY 166 ----- 166
Db 4603 GAAAGTGAAGAACTGTACACACAGACTGAACTGGGTGTGTCAAACTGAAAAAA 4662
QY 166 ----- 166
Db 4663 AAAAAAATCATTGAGATGAAAGAGATCAGCAAGCTGTAACATGGGCTATTTC 4722
QY 166 ----- 166
Db 4723 ATGTACAGATGTGATTTTACTGAACAATTTCTTCAAGATCTAGGCCCTGAAGACT 4782
QY 166 ----- 166
Db 4783 CACTGCTTATCTGTGAACATCTGAACCTGAAATGGATTTGTAGCTTTGTAGA 4842
QY 166 ----- 166
Db 4843 CAAAGTGAATTAACAACATCTGCACAAACCAAGCCCCCTTTCTGTTCCTA 4902
QY 167 --GlnArgGluLeuLeuHisIleLeuLeuAlaTyrGluIuTyrAsnPro----- 182
Db 4903 GGCAGCGGAACTACTCCACATCTCTCTGCAATGAGATTAACCCGGTGAATTC 4962
QY 182 ----- 182
Db 4963 CCGGAGTAGGTTCCCGGGCATATTTCATATTGACAGAGTGGGTGTGTGGGGG 5022
QY 182 ----- 182
Db 5023 TGTCTGTCTCTTTAAAGTAGTATTGTGACCAACAGATTAATAGAGGTAGATGT 5082
QY 182 ----- 182
Db 5083 CAGCTCACCGCTGGCATTAACCTCCAGAGAGGGGTGTCTCAAGGGGTCAAGCTGAGA 5142
QY 182 ----- 182
Db 5143 CACAAAGAGTACAGGCGCCGAGACTCTGTGTGTCACTGGGCTGACACCACTTTCAGA 5202
QY 182 ----- 182
Db 5203 ACAGAAATGACGCCCTCTCTCTGGGGTCCGCCAAAGCCAGAGCTTGGAGCATGCG 5262
QY 182 ----- 182

Db 5263 ACACAGATGTGTATACAGACATTTTGGACAAGGTGTGAAGTGTATGACTT 5322
QY 182 ----- 182
Db 5323 GGCCTTGTGATGAATAGATGTGATCTGTAGAGAGCCTCTTTTACAGAGAGCCTCT 5382
QY 182 ----- 182
Db 5383 CCTTACAGAGAGCCTCTCCAGTCACTCTGTCCCTCTCCAAATGATGATCTCCAGG 5442
QY 183 -----GluValG 185
Db 5443 TGACCTCAGCCCTCCAGGTGATGTCTTCCATGTGTACTGTGCTCTTGACAGAGTGG 5502
QY 185 |YTYTCyARgAspLeuSerHisIleAlaLeuPheLeuLeuTyrLeuProGluGlu 205
Db 5503 GCTACTGACGGGACCTGAGCCATGCGCGCTTGTCTCTCTATCTTCTCTGAGGAGG 5562
QY 205 sPAlaPheTrrPAlaLeuValGlnLeuLeuAlaSerGluArgHisSerLeuGln----- 222
Db 5563 ATGCATTTCTGGGCACTGTGTGACGTGTGTGCGCATGAGAGGCACTCCCTGCA-GGGTAAG 5621
QY 222 ----- 222
Db 5622 TGAACAGCTGCCCCGGGGACCTCTGACAGCAGACTGGGGATGGCCACTGCGCGGT 5681
QY 222 ----- 222
Db 5682 GATCAGACTTTCAGCCAAAGGACCCCTCTGTGTGCGCAAGTTTGGAGACTTTAGG 5741
QY 222 ----- 222
Db 5742 ATGTCTGTCTGAGGCTCCACAGAGATTCAGGCTGACCCCAAGCCCAATCAGAG 5801
QY 222 ----- 222
Db 5802 CCTCTATCCCATGACAGAGGAGATCATCTCTCCCGTGGCCACCTCTGTGTCTG 5861
QY 222 ----- 222
Db 5862 GAGCAAGCCCTCGGGCTGTATCTGTGACGTGACTCTCCCTCCCTGAGATCTCC 5921
QY 222 ----- 222
Db 5922 TGCCTCCAGTGCCTCGGCTCTCTGTGCATCGGTGCCAGAAATGGCCCAAGGCC 5981
QY 222 ----- 222
Db 5982 CAGGTGGACGATCTCCCATCCCTGTTCCCTGTGGCCGACCCCATACAGAGATGA 6041
QY 222 ----- 222
Db 6042 CCGGAGGCCAGCGCCCAACCATGTTCCGGCCGCTGTGTGAGCTGAAGTCAAGCTT 6101
QY 222 ----- 222
Db 6102 GCCCTTTTTCACCTTGCCCAAGAGGCTCCAGGGGAACCTTCAGACGACTCCAGGGA 6161
QY 222 ----- 222
Db 6162 ATGTTCGCGCCCACTCCCAAGGTAAAGCCGATGTTGGGGTCAACAGATGGAGGG 6221
QY 222 ----- 222
Db 6222 TGGAGTAGCCTTGGGGTTTGGGGCTCTCCAGTGCACACTTGTGAGCTGATGCT 6281
QY 223 -----GlyPheHis 226
Db 6282 CCACATCTTGGGGGAAGCTGTGATTTGATGATGGGCTGGGGGCTTCTCAGATTTTACA 6341
QY 226 ePProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnIleValValAlaIleHis 246

```

Db      6342 GCCCAATGCGGAGCGGTCCAGGGGCTCCAGACCAAGAGCATGTGTAGCCACGT 6401
Qy      246  |||||ProlyThrMetGly----- 252
Db      6402 CACATTCAGAGACCATGGGCGATCAGGTAGTTATGTCTCCCTCAGCTCTTCCAGAGG 6461
Qy      252 ----- 252
Db      6462 CCTGCTCCCGTGGGCTGTAGAGAGAGGGGGCTGGGGCCCTCTGTGGGGCTGTGAC 6521
Qy      252 ----- 252
Db      6522 TGGCTAGTCCAGCCAGGCGCTGACCTGGAGAGTCCGAGTCTTCATGGGCTGGAGTTG 6581
Qy      252 ----- 252
Db      6582 GTTTCCTTCTCTGCTCCCTGGAGGAGACAGAGGACAGAGATGGGGGCCAGCTCCCGAGA 6641
Qy      252 ----- 252
Db      6642 GCAGGGCAAGGCGAGTGTGTCCACCGGAGTGTGGAAAGGTGACAGTGTGTGGGAGC 6701
Qy      252 ----- 252
Db      6702 TCTGACACCGCCAGTGTCTGCACTAGGGAGAGGTCTTCAGAGACCCGAGAGAGG 6761
Qy      253 ----- 258
Db      6762 AGGTTTTAGGGAGCCAGCGCTGAGCCTCTGTGCTTCATCAGAGACAGAGAG 6821
Qy      258 sPLeuCySgLyGlnCySerProLeuGlyCySLeuIleArgIleLeuIleAsp-Gly-- 276
Db      6822 ATCTATGTGGGACAGTGTCTCCCGTTAGGCTGTCTCATCCGATATGATTGACGGGTAA 6881
Qy      276 ----- 276
Db      6882 GGAGGATAGGAGACCTGTGCTCAGAGACCTTCTTGCCCTGCAAGTCCCTGCTTCCC 6941
Qy      276 ----- 276
Db      6942 AGCCCGGGGGTCTGGCTCATCTCCAGCCCAAGAGGCTCAGCGGGTCCCAAGAGACA 7001
Qy      276 ----- 276
Db      7002 CACAGCAAAACCTCTGCCCAAGAGGGGTCAATCCAGGGCAATGGCTGGGCTCAGGCC 7061
Qy      276 ----- 276
Db      7062 CAGCTCATGGGAGACTGGGCGAGGACCCGACTTGAGAGGGCTCAGGGAAGCTCAAGC 7121
Qy      276 ----- 276
Db      7122 CTTGGGCAAGCCCTCTCTCCAGAGGACCATCCCACTCAATGATGTCCTCCCATGAG 7181
Qy      276 ----- 276
Db      7182 GAGCTTCAAGACCTTGTCTGACCCAGCGTCTTGAGGGCTCAGGCACTCATGGGAA 7241
Qy      276 ----- 276
Db      7242 GGTCACTGACTGTGAGACTGAAGCCCGAGTGTGCGAGCTGAGCCACAGCCCAAGC 7301
Qy      277 ----- 283
Db      7302 TGGAGGACAGAGTCTTTCACACTGTGTCTCCACAGATCTCTCGGGCTCACCTG 7361
Qy      284 ArgLeuTPAspValTyrLeuValGlnGlnIleLeuMetProIleThrArgIle 303
Db      7362 CGCCGTGGGAGCGTGTATCTGTAGAGGCGACAGCGCTTGATCCGATTAACAAGATC 7421
Qy      304 AlaPheLysValGlnGln----- 309
Db      7422 GCTTTAAGTTTCAAGCAGATGATCTAAGTGTGCTCCAGCGGGGCTTGGGAGACCTGGG 7481

```

```

Qy      309 ----- 309
Db      7482 GTACAGCCCGACTGGCCCGAGGGCAGTTCTCTACACTGTCTCATGATCCGCTTTCT 7541
Qy      309 ----- 309
Db      7542 GGCCAGAGGAGTCCGGCCAGGTGGCTGGGCGAGCACTGTGACACCGAGCCATCC 7601
Qy      309 ----- 309
Db      7602 CTCACATATCCGAGCAGGGAAGTCTCACCACACTTCGACTTTCATCTGGGTCCCGAG 7661
Qy      309 ----- 309
Db      7662 CCACAGTCTCTGTGTATATCTGACACTGGGGTGGCCACAAAGATCCGACCGAC 7721
Qy      309 ----- 309
Db      7722 CAGTAGAGACTGAAGTGGCCACGGGGATGAGCTGTGACCATTCAGGTAACCTCCCT 7781
Qy      310 ----- 321
Db      7782 GGCCTGATATCACCTCTGTCTCTAGAGGCTTCAAGAGAGTCCAGGTGTGGCCGTGG 7841
Qy      322 AlaArgPheCySAsnArgPheValAspThrTyrAlaArgAspGluAspThrValLeuLys 341
Db      7842 GACGGTTTTTGCAACCGGTTCTGTGATACCTGGGCGAGGATGAGCACTGTGCTCAAG 7901
Qy      342 HisLeuArgLAspMetLysLysLeuThrArgLysGlnGlyAspLeuProProAla 361
Db      7902 CATCTTAGGGCTCTATGAAAGAACTAACAGAAAGCAGGGGACCTGCCACCCACAGG- 7960
Qy      361 ----- 361
Db      7961 TGGGCTCAGTGCATGTCCCTCCCATGTCACCTCTGGGGTACTGATGAGGAGG 8020
Qy      361 ----- 361
Db      8021 TGCCCGGAGCCGCAACCTACTACCTGGGCTTCTCTTCACTTTCTTCTCTCTT 8080
Qy      361 ----- 361
Db      8081 CCTCTGAACTTAAGAAAGTACAGAGGCCCAACCGGTCTTCAGGGCAGGCTCATGTC 8140
Qy      361 ----- 361
Db      8141 GTGTATATCGACATGTGTGACGAGAGGGGATGTGGCAAGACCTTCAACAAGC 8200
Qy      361 ----- 361
Db      8201 CCCCTCCACTTTCACAGGTGTCTTCCCTCTCTCCCTGCGAGGGCTTCAAGTTACTAG 8260
Qy      361 ----- 361
Db      8261 CGAGCCAGACCATTTGTGGAGACCCCGCCCTCTCTGGAAGACCAACGAGCTCAGA 8320
Qy      361 ----- 361
Db      8321 GAGCAGCAGAGCCCTCACTCTGCAAGCTCTCTCAAGTTGCCAGACAGAAAGCCTG 8380
Qy      361 ----- 361
Db      8381 GAGCCAGGAGACAAAGGAATCCGTGTCTCTGACCAAGAGCATTCAGGGAGAGGGCCC 8440
Qy      361 ----- 361
Db      8441 AGAGCAGAGCAAGATTCAGCCAGAAAGTGGAAAGGTCAAGTCTGAGATGAGACTGGGC 8500
Qy      361 ----- 361
Db      8501 AGCCAGAGGAGCAGAGGTGACCAAGTCCGGGCCCAATCAACCACTGCGGAGACGGGT 8560

```

```

QY 361 ----- 361
DB 8561 CCCACGTGAGTGACAAAGGGCTGGGTGACATCCAAAGGCCCTCCCACTGATGTTCTGA 8620
QY 361 ----- 361
DB 8621 CTGGGGCCGTATCCCAAGGCCCAAGAGCCCTGGAGCAAGATGTGTGTGAGGAAGAGCCCC 8680
QY 361 ----- 361
DB 8681 AGCCAGTCTGAACCTGGGGGAGATCCAGAGCAACCCGCAATGCCAGACAGCTTCCC 8740
QY 361 ----- 361
DB 8741 CACGCCAGGACATGACACCCCTCTCTGGATCAGAGACTACAGGCGTGTCTCGGT 8800
QY 361 ----- 361
DB 8801 GTCAAGCCACGGGGGCAACAGAGGACCCGAGACTCCGAGATGCAAGAGTGGGGCC 8860
QY 361 ----- 361
DB 8861 CAGCCGGAGAGGCTGCTGGGCTCACTGAGATGCTGACCGCGTCTGTTTCTTTCA 8920
QY 362 ---LysProGluGlnGlySerSerAlaSerProValProAlaSerArgGlyLys 380
DB 8921 GCCAAACCCGACAGAGGCTGTCTGCGATCAGAGCCCTGTCCGCTTCACTGCGGAG 8980
QY 381 ThrLeuCybGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro 400
DB 8981 ACCCTCTGCAAGGGGAGCAGAGGAGCCCTCCAGGCCCAAGCCGGTTCGCGGGCC 9040
QY 401 IleTyrSerAlaSerProProArgAlaProArgSerSerThrProCybProGlyValAla 420
DB 9041 ATTGTGTACGCTTCCCGCCAGCGGCACTCTTCTTCCACACCCGTCTGTGGGCT 9100
QY 421 ValArgGluAspThrTyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGln 440
DB 9101 GTCCGGAGAGACCTTACCTCTGTGGCATTCAGAGGTGTCCAGCCCGGCTTGGCTCAG 9160
QY 441 GlyGlyProGlnGlySerTyrPargPheLeuGlnTyrPasnSerMetProArgLeuProThr 460
DB 9161 GGAGGACCTCAAGGTTCCGTGAGATTCTCGAGGTGAATCCATGCGCCGCTCCCAAG 9220
QY 461 AspLeuAspValGlyGlyProTyrPheArgHisTyrAspPheArgGlnSerCybTyrVal 480
DB 9221 GACCTGAGAGTGAAGGCGCTTGTGTTCCGCAATTATGATTGACAGAGCTGTGGGCTC 9280
QY 481 ArgAlaIleSerGlnGluAspGlnLeuAlaProCybTyrGlnAlaGlnIleAspProAlaGln 500
DB 9281 CGTGCAATATCCAGAGAGACAGAGCTGGGCCCTGTGCGAGCTGAACACCCCTGCGAG 9340
QY 501 ArgValArgSerSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg 520
DB 9341 CGGGGAGATGCGCTTGTGCTGACACCACTGATTCGACCAAGGAGCCCTTCCACA 9400
QY 521 AlaArgAspGlnGlnProCybAlaProThrSerGlyProCybLeuCybGlyLeuHisLeu 540
DB 9401 GCTAGAGAGCAAGACAGAGTGTCTCCCACTCAGAGGCGCTTGTGCGCTTCCACTTG 9460
QY 541 GluSerSerGlnPheProProGlyPhe 549
DB 9461 GAAAGTTCTCAAGTTCCCTCCAGGCTTC 9487

```

```

; CURRENT FILLING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1824
; LENGTH: 9805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1824

Alignment Scores:
Pred. No.: 1,586-128 Length: 9805
Score: 1507.50 Matches: 487
Percent Similarity: 20.97% Conservative: 6
Best Local Similarity: 20.71% Mismatches: 17
Query Match: 50.54% Indels: 1844
DB: 14 Gaps: 11

US-10-071-838-2 (1-549) x US-10-091-504-1824 (1-9805)

QY 24 LysGlyHisArgAlaGlyLeuProGluAspLysGly-----ProLysProPheArgSer 41
DB 2492 CAGGCTCCAGAGCCCAAGACCCAGACATCCGCGCGGCTGTGGAGAGCTGGAGCTCC 2551
QY 42 TyrAsnAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProProLeuThr 61
DB 2552 GCTAATCTCAACATCCCTCATTTG----- 2575
QY 62 AlaArgGluAlaLys-GlnIleArgGluIleSerArgLysSerLysTyrValAspPhe 81
DB 2576 -----ACAGCAAAATTCGCGGAGATCAGCCGAAAGACCAAGTGGTGTAT 2623
QY 81 IleuGlyAspTyrGlyLysTyrLysSerSerArgLys----- 93
DB 2624 GCTGGAGACTGGAGAAATACAAAGACGAAAGTAACTGACGTGAGAGAGAGCA 2683
QY 93 ----- 93
DB 2684 CTCTCTGACAGACAGGGGACAGGACCCATGCTTGTGCTTGGACATCAGCTCTC 2743
QY 93 ----- 93
DB 2744 AGAGGGTGGGGGCACTGTCTCGCCAGAGACTGCAAGGCTGTGCGCAGATTTC 2803
QY 93 ----- 93
DB 2804 TGCCTATTCTGCAAGCGTCACTTTCAGAGAGGAATCTGAATCTAGGCTGGAATAC 2863
QY 93 ----- 93
DB 2864 CCGAGCTCAAGGCTAAGGATGCTGCTGATCCTGAAGAGAAAGTTCAATCAGA 2923
QY 93 ----- 93
DB 2924 GTTTCAGCTGAGTGTCACTCACTTTTCACTCTGGAGAGAGACCTGTCCAGC 2983
QY 93 ----- 93
DB 2984 TTGATCTACCTCTAAGAGATGAGGCAAAACCAACAAATTTCCAGATCCCGG 3043
QY 93 ----- 93
DB 3044 GCTTGTGCTCACTAGGGGTACCCCGTGGCTGTGACACCAAGATTGTTTCTGCCACA 3103
QY 94 -LeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyProMetTyrSerVal 113
DB 3104 GCTCATATGATCGAGATACAAAGGAATGCCATGAATCCGAGGCGCCGATGTGTAGT 3163
QY 113 IleuLeuAsnIleGluGluMetLysLeuLysAsnProGlyArgGlyGlnIle----- 130
DB 3164 CTTCTTGAACATTTAGAAATGAAATGAAAAACCCGAGATACCAAGT-ACGCTCAG 3222
QY 130 ----- 130

```

Db	4303	GGAAAGCCCACTGAGAAAAGGAAACGAGCCGGTGAATGCCGACGAAACGTGGGTGATCC	4362
OY	166	-----	166
Db	4363	TAGATGCAATTTTGTGTGAGGGACAGAGCCAGACCCAAATAGCTACAGTAGATGATCC	4422
OY	166	-----	166
Db	4423	ATTCTAGGCCAATTTGTGAAAAAGGCCAAACACAGGGACTGAGAAAGCAGTCTGGGTGCC	4482
OY	166	-----	166
Db	4483	AGGGGCTGACGGATCCGGGGAGAGGCTGGTGCAATAGGGCCACCTTGAGACTTGGAGA	4542
OY	166	-----	166
Db	4543	TGAAGAGTGCCTCCAGAGGGGCTGAGCGGTGGCCGGGAGACTCTGCAATTTGGTTTG	4602
OY	166	-----	166
Db	4603	GAACCGTGAGAACTGTACACACACAGACTGAACTGGCGTGTGCAAACTGAAAAAA	4662
OY	166	-----	166
Db	4663	AAAAAAATCATTCAGAGTGAAGAAAGATCAGGCAAGTCACTGTACAATCGGCTATTGC	4722
OY	166	-----	166
Db	4723	ATGTCAAGATGTGATTTTACTGAAACATTTCTTCAAGTCTCAGGCCCTGAAAGCT	4782
OY	166	-----	166
Db	4783	CACCTGTATCTGTGAACATCTGAACCTGAAATGGGATTTGCTGTAGCTTGTGAGA	4842
OY	166	-----	166
Db	4843	CAAAAGTAATTAACACATCTGCACAAAACAACAAAGCCCTTCTCTGTTCCGA	4902
OY	167	--GhaTgSluleuLeuHs11eLeuLeuA1aTyrgluGurYzAnPro-----182	
Db	4903	GGCAGCGGGAACTACTCCACATCCTCTGTGCATATAGAGATTAACCCGGTGAATTC	4962
OY	182	-----	182
Db	4963	CCGGCAGTGAAGTTCCGGGCGCATATTTCATATTTGACAGAGTGGTGTCTGAGGGGG	5022
OY	182	-----	182
Db	5023	TGTCTGTCTCTTTTAAGTTAGTATTTGTGACCAACAGATATAGAGTAGATGT	5082
OY	182	-----	182
Db	5083	CAGCTCACCGTGCAATMAACTCAAGAMAGGGGTGTCTCAAGGGGTCAAGCTGAGA	5142
OY	182	-----	182
Db	5143	CACAAAGAGTCAAGGCGCGGACTCTGTGTTCACCTGGGCTGACCACTTTCAGA	5202
OY	182	-----	182
Db	5203	ACAGAAATGACGCTCTCTCTGGGGCTGCCCAAGCCAGGAGCTTGGCAGATCCG	5262
OY	182	-----	182
Db	5263	ACAAGAGTGTGTATACGACACATTTTGACAAAGTGTGAAGTGCCTGATGACTT	5322
OY	182	-----	182
Db	5323	GGCTTTGTCAATGAATGATGTGATCTGAGGAACCTCTTTTCAAGAGAACCTTCT	5382
OY	182	-----	182
Db	5383	CTTTACAGGAAGCTCTTCCACGTACCTTGTCCCTTCACATGACATGATGCTCCAGG	5442

Oy	183		-----gluVal ¹⁶ G	185
Db	5443	TGACCTCAGCCCTCCCAGATGATGTCTTCCATTGGTAGACTCGGCTTCGCAGAGAGGTG		5502
Oy	185	IYYTYAARPAAPLeuSerHisIleAlaLeuPheLeuNylrLeuProIugIua		205
Db	5503	GCTACTGCAGGGAACTTGAGCCAACAATGCCGCTTTCTCTCATCTTCCCTGAGAGG		5562
Oy	205	sPaLaPhetPalaleuValGlnleuleuAlaserGluArgHiserLeuIn-----		222
Db	5563	ATGCATTTGGGCACTGGGTGACGTGTGGCCAIGTAAGGCACTCCCTGA-GGSTAAg		5621
Oy	222	-----		222
Db	5622	TGAACAGCTGCCCGGGAGCCTCCTGACAGCACACTTGGGANTGGCCA-CCTGCGGGGT		5681
Oy	222	-----		222
Db	5682	GATCACAAGCTTTCAGACCAGAAGCAACCTCTGTGTGCGCAGCTGTGTGGAGACTTTAG		5741
Oy	222	-----		222
Db	5742	ATGTCTCTGTGAGGGTCCCAAGAGATCCAGSGCTGACCCCCCAAAGCCAAATCAGACG		5801
Oy	222	-----		222
Db	5802	CCTCTCATCCCATCAGAGAGGAGCATCTACCTCCCCGTGGCCACCTCTGTGTCTG		5861
Oy	222	-----		222
Db	5862	GAGCCAGCCTCCGGCTCTGATYCTGTGACAGTGACTTCCCTCCCTGAGAGTCTCC		5921
Oy	222	-----		222
Db	5922	TGCCTCCAGCTGCCCGGGCTCTGTCTCCATCGGTGCCAAGAATGGGCGGACCAAGCC		5981
Oy	222	-----		222
Db	5982	CAGGTGCAGAGATCTCCCATCTCCCTGTCTCCCTGGCCGACCCCACTAACAGAGATGA		6041
Oy	222	-----		222
Db	6042	CCGGAGAGCCAGGCGCCACCACTTCGGCGCGCTGTCTGTGGCTGAAGTCAAGCTT		6101
Oy	222	-----		222
Db	6102	GCCTTTTTTGACCTTGSCCGAGAGGCTCCAGGGGAACCTCAGCGAGCTCCAGGGA		6161
Oy	222	-----		222
Db	6162	ATGTTCGGCCCACTCCCAAGGTAAAGCCGCAATGTGGGTGACACAGATGGAGGG		6221
Oy	222	-----		222
Db	6222	TGGGAGTAGCCTTGGGGTTTGGGGGCTCTCCAGCTGCCAAGCTTTCAGACTATGCT		6281
Oy	223	-----glyPheHis	226	
Db	6282	CCACATCTTGGGGAGAGGCTCTGATTCATGATGGCTGGGGGCTTCCACAGAAVTTTACA		6341
Oy	226	ePrOanGIvGIThyrValGlnGlyLeuGlnaSPGLnGlnIsvValalaThrS		246
Db	6342	GCCCAATGGGGGAGCCTCCAGGGGCTCCAGAACCAAGAGCATGTGGTtagcAcGT		6401
Oy	246	erGlnProLyserThrMetGly-----		252
Db	6402	CACATTCAMAGACATGGGCGATCAGTGAATTATGTCCCTTCAGCTTTCACAGAG		6461
Oy	252	-----		252
Db	6462	CCCTGCTCCCGTGGGCTGTAGAGACAGGGGGGCTGGGGCCCTCGTGGGGCTGTGAC		6521

[illegible]

Db 7602 CTCACATGATCCGAGACAGGGAAGTCTACCACTCTCGACTTTTCATCTGGTCCCGAG 7661
 QY 309 ----- 309
 Db 7662 CCACAGTCTCTGTGTATATCTGACACTGGGGTGGCCAAAGATCCGGACCGAC 7721
 QY 309 ----- 309
 Db 7722 CAGTAGAGACTGAAGTGGCCACGGGGTATGAGCTGTGACCATTTCCAGTACTCCCT 7781
 QY 310 ----- 7781
 Db 7782 GGCCTGATATCCACCTGCTCCCTAGAGCGCTCAGAGAGCGTCCAGGTGTGGCCGGTGG 7841
 QY 322 A1aArgPheCyAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeuLys 7841
 Db 7842 GCACGTTTTTGCAACCGGTTCTGTGATACCTGGGCGAGGATGAGGACACTGTGCTCAAG 7901
 QY 342 HisLeuArgLAspSerThrLysLysLysLysLysLysLysLysLysLysLysLysLys 7901
 Db 7902 CATCTTAGGGCTCTGTATGAGAACTAACAGAAAGAGGGGGAACCTGCCACCCCGAGG- 7960
 QY 361 ----- 7960
 Db 7961 TGGGCTCCAGTGCATGTCCCTCCCATGTCACTCTGGGGTATGACTAGTAGGGAG 8020
 QY 361 ----- 8020
 Db 8021 TGGCCGGAGCCCGCAACCTTACTACCTGAGGCTTCTCTTCACTTTTCTTCTCTT 8080
 QY 361 ----- 8080
 Db 8081 CCTCTGACTCTTAGAAAGTACAGAGAGCCCAACGGTCTTCAAGGAGCGCTCAGTGC 8140
 QY 361 ----- 8140
 Db 8141 GTGTACTGACATGTCTGTGACGACAGAGGGGATGTGGGCAAGACCTCCACAAAGC 8200
 QY 361 ----- 8200
 Db 8201 CCCCTCCACTTTCCAGGGTGTCTCCTCTTCCCTCGAGGGCCTCCAGTTACTAGA 8260
 QY 361 ----- 8260
 Db 8261 CGAGCCAGACCATTTTGTGTGAGAGCCCGCCTCTCCAGAAAGCAACCAAGCTCAGA 8320
 QY 361 ----- 8320
 Db 8321 GAGCAGCAGAGCCCTCACTCCTGACGCTCCTCCAGGTTTCCAGAGCAAGAGCTTG 8380
 QY 361 ----- 8380
 Db 8381 GAGCCAGGAGACAGAGGAATCCTGTCTCCCTGACCAAGAGCATTTCAAGGAGAGGGCCC 8440
 QY 361 ----- 8440
 Db 8441 AGAGCCAGAGCCAAAGATTCAAGCCAAAGTGGAAAGGTCACTCTGATGATGAGTGGGC 8500
 QY 361 ----- 8500
 Db 8501 AGCCCAAGAGGAGGAGGAGTACCACTCCGGGCCCAATCACCACACTGGGAGAGCGGGT 8560
 QY 361 ----- 8560
 Db 8561 CCCCACTGAGGTGACAAAGGGCTGGGTGATCCAAAGGCCCTCCCACTGAGTTCTGA 8620
 QY 361 ----- 8620
 Db 8621 CTGGGGCCGTATCCAGGCCCAGAGCCCTGGAGCAGAGGTGTGTGGCAGAGAGCCCCC 8680
 QY 361 ----- 8680

Db 8681 AGCCAGTGAACCTTGGGGGAGTCCAGAGAGCCACCCGCCATGACAGACGCTTCCC 8740
 QY 361 ----- 8740
 Db 8741 CAGCCAGGAGATGACCCCTCTCTGGGATGACAGACTACAGCGTGTCTCGGT 8800
 QY 361 ----- 8800
 Db 8801 GTACAGCACGAGGGGCCACACAGAGGACCCGAGAGCTCCAGATTCAGGACAGTGGGCC 8860
 QY 361 ----- 8860
 Db 8861 CAGCCGGAAGAGCTGCGTGGGCTCACTGAGATGTGACCGGCTGTTTCTTTCA 8920
 QY 362 ---LysPProGluGluGlySerSerLAspArgProValProLAspArgGlyLys 8920
 Db 8921 GCCAAACCCGAGCAAGGTCGTGGCATTCAGGCTGTGCCGCTTCACTGTGGGGAG 8980
 QY 381 ThrLeuCyLysGlyAspArgGluAlaProProGlyProProAlaArgPheProArgPro 400
 Db 8981 ACCCTCTGAGAGGGGAGACAGAGAGCCCTCAGAGCCACAGCCGCTTCCCGGCC 9040
 QY 401 IleTProSerLAspProProArgAlaProArgSerSerThrProCysProGlyValAla 420
 Db 9041 ATTGTGTAGCTTCCCGCCACGCGGACCTGTTCTTCCACACCTGTCTGTGGGGCT 9100
 QY 421 ValArgLAspThrTrpProValGlyThrGluGlyValProSerProAlaLeuAlaGlu 440
 Db 9101 GTCCGGAGAGCACCTACCTGTGGGACATCAAGGGTGTGCCAGCCCGGCTGGGCTCAG 9160
 QY 441 GlyGlyProGluGlySerTrpArgPheLeuGluTrpAsnSerMetProArgLeuProThr 460
 Db 9161 GAGAGACCTCAGAGGTTCTGTGAGATTCTGAGTGAATCTCATGCCCGCTCCCAAG 9220
 QY 461 AspLeuAspValGluGlyProTrpPheArgHisTrpAspPheArgLysCysTrpVal 480
 Db 9221 GACTGGAGCTAAGGGCCTTGTGTTCGCAATATGATTTCAGACAGAGCTGTGGGTC 9280
 QY 481 ArgAlaLysSerGluAspGluLeuAlaProCysTrpGluAlaGluHisProAlaGlu 500
 Db 9281 CGGCAATATCCAGAGGAGACAGCTGGCCCTCTGAGAGCTGAAACACCTTGGGAG 9340
 QY 501 ArgValArgSerLAspAlaAlaProSerThrAspSerAspGluGlyThrProPheArg 520
 Db 9341 CGGTTAGATCGGCTTGTGTGTGACACCACTGATTCGACAGGGCACCCCTTCAGA 9400
 QY 521 AlaArgAspGluGlnProCysAlaProThrSerGlyProCysLeuCyGlyLeuHisLeu 540
 Db 9401 GCTAGGAGAGAACACAGTGTGTCTCCCACTCAGGGCTTGTCTGTGGGCTTCCACTTG 9460
 QY 541 GluSerSerGlnPheProProGlyPhe 549
 Db 9461 GAAAGTCTCAGTTCTCTCCAGGCTTC 9487

Search completed: February 4, 2005, 13:50:57
 Job time : 991 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 4, 2005, 08:05:19 ; Search time 136 Seconds

(without alignments)
2869.287 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983

Sequence: 1 MDVBEVAGSWMAQEREDIM.....TSGPCLGHLHSSQPPPGF 549

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO_epool/US10071838/runat_03022005_071219_215/app_query.fasta_1.711
-DB=Issued Patents NA -QFMT=fastcap -SUFFIX=ext1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPRMT=pco -NOR=ext -HEARSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10071838@cgn2_1.1.69,@runat_03022005_071219_215 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2850	95.5	2084	4	US-09-799-451-692
2	2827	94.8	2072	4	US-09-799-451-691
3	2133.5	71.5	8201	1	US-08-253-155A-9
4	344	11.5	1981	4	US-09-620-312D-715
5	337	11.3	1823	4	US-09-774-528-336
6	258	8.6	3039	4	US-09-620-312D-675
7	253	8.5	2955	4	US-09-620-312D-676
8	238	8.0	1083	4	US-09-248-796A-5457
9	222	7.4	1887	4	US-09-270-767-12482
10	194.5	6.5	806	4	US-09-270-767-9899
11	182.5	6.1	4039	1	US-08-363-300-1
12	172	5.8	1056	4	US-09-248-796A-1133

13	154	5.2	1659	4	US-09-248-796A-1820	Sequence 1820, Ap
14	149	5.0	7527	4	US-09-252-991A-71	Sequence 71, Ap1
15	148.5	5.0	3763	4	US-09-919-039-243	Sequence 243, Ap1
16	148.5	5.0	5185	4	US-09-976-599-640	Sequence 640, Ap1
17	147.5	4.9	723	4	US-09-270-767-1491	Sequence 1491, Ap
18	147.5	4.9	723	4	US-09-270-767-1491	Sequence 16773, A
19	146.5	4.9	50937	3	US-09-428-517-1	Sequence 1, Ap11
20	146	4.9	3204	4	US-09-799-451-340	Sequence 340, Ap1
21	146	4.9	3102	4	US-09-252-991A-4429	Sequence 4429, Ap
22	146	4.9	43280	2	US-08-804-227C-1	Sequence 1, Ap11
23	145.5	4.9	1368	4	US-09-252-991A-64	Sequence 64, Ap1
24	145.5	4.9	4403765	3	US-09-103-840A-2	Sequence 2, Ap11
25	145	4.9	4411529	3	US-09-103-840A-1	Sequence 1, Ap11
26	144.5	4.8	2009	4	US-09-646-403-1	Sequence 74, Ap1
27	144.5	4.8	12249	4	US-09-266-965-74	Sequence 96, Ap1
28	144.5	4.8	18331	4	US-09-266-965-96	Sequence 207, Ap1
29	143	4.8	1421	3	US-08-943-731-207	Sequence 22, Ap11
30	143	4.8	2363	4	US-09-818-780-22	Sequence 4, Ap11
31	143	4.8	6085	3	US-09-029-603-4	Sequence 5, Ap11
32	143	4.8	20084	3	US-08-943-731-5	Sequence 18, Ap11
33	142.5	4.8	990	4	US-09-266-965-18	Sequence 14376, A
34	142.5	4.8	1593	4	US-09-252-991A-14376	Sequence 4632, Ap
35	142	4.8	1788	4	US-09-252-991A-632	Sequence 4181, Ap
36	142	4.8	5184	4	US-09-845-583A-9	Sequence 4, Ap11
37	142	4.8	5184	4	US-09-561-709B-4	Sequence 209, Ap1
38	142	4.8	8037	4	US-09-774-528-29	Sequence 1, Ap11
39	142	4.8	8037	4	US-09-292-034-1	Sequence 486, Ap1
40	141	4.7	34094	4	US-09-620-405B-486	Sequence 486, Ap1
41	140	4.7	827	4	US-09-604-287A-486	Sequence 486, Ap1
42	140	4.7	827	4	US-09-834-759-486	Sequence 484, Ap1
43	140	4.7	1518	4	US-09-620-405B-484	Sequence 484, Ap1
44	140	4.7	1518	4	US-09-604-287A-484	Sequence 484, Ap1
45	140	4.7	1518	4	US-09-604-287A-484	Sequence 484, Ap1

ALIGNMENTS

RESULT 1

US-09-799-451-692

Sequence 692, Application US/09799451

Patent No. 6783969

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Zhou, Ping

APPLICANT: Goodrich, Ryle

APPLICANT: Asundi, Vinod

APPLICANT: Ren, Feiyang

APPLICANT: Zhang, Jie

APPLICANT: Xue, Aidong J.

APPLICANT: Zhao, Qing A.

APPLICANT: Wang, Qian-Rui

APPLICANT: Ma, Yungling

APPLICANT: Yamazaki, Victoria

APPLICANT: Chen, Rui-hong

APPLICANT: Wang, Zhiwei

APPLICANT: Yang, Dunhui

APPLICANT: Yang, Yonghong

APPLICANT: Wehtman, Tom

APPLICANT: Ghosh, Reena

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 803

CURRENT APPLICATION NUMBER: US/09/799,451

CURRENT FILING DATE: 2001-03-05

NUMBER OF SEQ ID NOS: 948

SOFTWARE: pc FL_genes Version 2.0

SEQ ID NO: 692

LENGTH: 2084

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS


```

1  APPLICANT: Ghosh, Reena
2  APPLICANT: Dmanac, Rajee T.
3  TITLE OF INVENTION: No. 6783936el Nucleic Acids and
4  TITLE OF INVENTION: Polypeptides
5  FILE REFERENCE: 803
6  CURRENT APPLICATION NUMBER: US/09/799,451
7  CURRENT FILING DATE: 2001-03-05
8  NUMBER OF SEQ ID NOS: 948
9  SOFTWARE: pt_fl_genes Version 2.0
10 SEQ ID NO 651
11
12 LENGTH: 2072
13
14 TYPE: DNA
15
16 ORGANISM: Homo sapiens
17
18 FEATURE:
19
20 NAME/KEY: CDS
21 LOCATION: (163)..(1743)
22
23 OS-09-799-451-691

```

Alignment Scores:	
Pred. No.:	1,11e-199
Score:	2837.00
Percent Similarity:	95.634
Best Local Similarity:	95.634
Query Match:	94.778
DB:	4
	Gaps: 1

US-10-071-838-2 (1-549) X US-09-799-451-691 (1-2072)

Qy	1	MeAspValValGluValAlaGlySerTPPTPrAlaGlnGluValAspLeu161	MeC	20
Db	163	ATGGACCTGGTAAGGATCGGGGCACTTGTGGGACAAAGCCGAGAGACATCAATTATG		222
Qy	21	LybTyGluLysGluYhiIsaTgaIagIleuProGluAspLysGlyProIysProPheArg		40
Db	223	AAATACGAAAGGAGCACCGAGCTGGCTGCCAGAGGACAAAGGGCTTAAGCTTTTCSA		282
Qy	41	SeTyrTrpAsnAsnValAspHisLeuGlyIleValHisGluThrGluLeuProLeu		60
Db	283	AGCTACAAACAACACCTGCATCTTTGGGGATTGATCAAGACCGAAGCTCTCTCTCG		342
Qy	61	ThrAlaArgGluAlaLysGlnIleArgTArgGluIleSerArgLysSerLysTyrValAsp		80
Db	343	ACTGCGGGGAGCGGAGGCAATTCGGCGGGAGATCAGCCGAAAGACAAATGGGTGGAT		402
Qy	81	MetIeuGluYAspTrpGluLysTyrLysSerSerArgLysLeuIleAspArgAlaTyrLys		100
Db	403	ATGCTGGGAGACTGGGAGAAATGCAAAACACAAAGACTCATGATCGAGCTAACAG		462
Qy	101	GlyMetProMetAsnIleArgGlyProMetTrpSerValIleuLeuAsnIleGluGlyMet		122
Db	463	GGAAATGCCATGAACATCCGGGGCCGAGATGGTCAATCTCTTAACACTGAAGAAATG		522
Qy	121	LybLeuLysAsnProGluYArgTyrGlnIleMetLysGluLysLysArgSerSerGlu		140
Db	523	AAGTTGAAAAACCCCGGAGAGATGCCAGATCATGAAGAGAAAGGCAAGAGCTCATCGAG		582
Qy	141	HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIleAspPheArg		162
Db	583	CACATCCAGGCAATCCACCGGACGTAAACCGGAGCACTTAAGAACATATATCTTCACG		642
Qy	161	AspArgTyrGlyThrLysGlnArgGluLysLeuHisIleLeuLeuAlaTyrGluGluTyr		180
Db	643	GATGTAAAGGAAACCAAGCAGGGGAACTAATCTCCACATCTCTCTGGCATTAAGAGAGAT		702
Qy	181	AsnProGluValGluTyrCysArgAspLeuSerHisIleAlaIleLeuPheLeuLeuTyr		202
Db	703	AACCCGGAGGTGGGCTAATGACAGGACCTGAAGCAACATGCGCGCTTCTCTCTCAT		762
Qy	201	LeuProGluGluAspAlaPheTPrAlaLeuValGlnLeuLeuAlaSerGluAspHisSer		222
Db	763	CTTCTCGAGGAGGATGATTTCTGGGCACTGGTGCACACTCTGGCCAGTGAAGGACATCC		822
Qy	221	LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGlnGlu		240

Db	823	CTGCAGGGGATTTCACAGCCCAAAATGGCGGAGCCGTCACGAGGCGCTCCAGAACCAACGAGAG	882
Qy	241	HisValValAlaThrSerGlnProIysThrMetGlyHisGlnAspIleValAspIleuCyS	260
Db	883	CATCTGGTAGCGACGTCACAAACCAAGCCATGGGCGATCAG-----	924
Qy	261	GlyGlnCysSerProIleuGlyCysLeuIleArgIleLeuIleAspGlyIleSerIleuGly	280
Db	925	-----ATCTCTCTCGGG	936
Qy	281	LeuThrIleuArgLeuTrpAspValTyrLeuValGlyGlnGlnAlaIleuMetProIle	300
Db	937	CTCACCCCTCGCCCTGTGGGAGCGTGTATCTGGTTAGAAAGCGCAACAGCGCTTGATGCCGATA	996
Qy	301	ThrArgIleAlaPheIysValGlnGlnIlyAspArgLeuThrIlyrThrSerArgGlyGlyPro	320
Db	997	ACAAAGATCGCCTTTAAGTTTCACAGAAAGCGCCTCACAGAAACCTCCAGAGTGTTGGCCG	1055
Qy	321	TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGlnAspThrValIleu	340
Db	1057	TGGCAGCGTTTTCGAACCGGTCGTGTTGATACCTGGGCGCAGGGAATGAGGACACTGAGCTC	1111
Qy	341	IlyHisIleuArgAlaSerMetIlyValLeuThrArgIlyGlnGlyAspIleuProTrpPro	360
Db	1117	AAGCATCTTAGAGGCGCTCTATGAAAMAATTAACAAGAAAGCAGGGGAGCACTGCACCCCA	1177
Qy	361	AlaIlyrProGlnGlnIlySerSerAlaSerArgProValProAlaSerArgIlyGlyIlys	380
Db	1177	GCCAAACCCGAGCAAGGTCCTCGGCATCCAGGCTGTGCGGCTTCAGTGGCGGAGAG	1233
Qy	381	ThrIleuCyIlyGlyAspArgGlnAlaProProGlyProProAlaArgPheProArgPro	400
Db	1237	ACCTCTGCAGAGGGGACAGGCAAGGCCCTCCAGGCCACCAAGCCCGGTTCCCGGGGCC	1299
Qy	401	IleTrpSerAlaSerProProArgAlaProArgSerSerThrProCysProGlyIlyAla	420
Db	1297	ATTGTGTGAGTTCCTCCGCGCACGGGCACTCGTCTTTCACACCTGTCTGTGGGGCT	1355
Qy	421	ValArgGlnAspThrTyrProValGlyThrGlnGlyValProSerProAlaIleuAlaGln	440
Db	1357	GTCGGGAGAGCACTACCTACCTGTGGGACCTCAGGGTGTGCCAGCCCGCCCTGGCTCAG	1411
Qy	441	GlyIlyrProGlnGlySerTrpArgPheLeuGlnTrpAsnSerMetProArgLeuProThr	460
Db	1417	GGAGAGACCTCAGGGGTTCTGTGAAGATCTCTGACAGTGAACCTCATGCCCCGCTCCCAAG	1477
Qy	461	AspIleuAspValGlyGlyProTrpPheArgHisTyrAspPheArgGlnSerCysTrpVal	480
Db	1477	GACCTGGAAGTGAAGGGGCCCTTGGTTCGCCCATTTATGATTTTCAGACAGCTGCTGGGCTC	1533
Qy	481	ArgAlaIleSerGlnIlyAspGlnIleuAlaProCysTrpGlnAlaGlnHisProAlaGln	500
Db	1537	CGTCCCATATATCCGAGAGAGACCACTAGGCCCTCCGTGCGAGGCGTGAACACCTTCGCGAG	1599
Qy	501	ArgValArgSerAlaPheAlaAlaProSerThrAspSerAspGlnGlyThrProPheArg	520
Db	1597	CGGGTGAATCGGCTTTCGCTGCACCCACGACCTGATTCGACACAGGAGCACCCCTTCACA	1655
Qy	521	AlaArgAspGlnGlnProCysAlaProThrSerGlyProCysIleuCyGlyIleuHisIleu	540
Db	1657	GCTAGGAGCGAAACGCGGTGTGCTCCACCTCAAGGCGCTTGCGCTCCACCTTG	1711
Qy	541	GluSerSerGlnPheProProGlyPhe	549
Db	1717	GAAGTTCCTAGTTCCTCCAGGGCTTC	1743

```

RESULT 3
US-08-253-155A-9
; Sequence 9, Application US/08253155A
; Patent No. 5691147
; GENERAL INFORMATION:
; APPLICANT: Gyuris, Jeno

```

APPLICANT: Draetta, Giulio
 TITLE OF INVENTION: CDK4 Binding Proteins
 NUMBER OF SEQUENCES: 95
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII(text)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/253,155A
 FILING DATE: 02-JUN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: MII-028
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-5941
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8201 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-253-155A-9

Alignment Scores:
 Pred. No.: 1.35e-147 Length: 8201
 Score: 2133.50 Matches: 408
 Percent Similarity: 77.23% Conservative: 33
 Best Local Similarity: 71.45% Mismatches: 58
 Query Match: 71.52% Indels: 72
 Gaps: 2

US-10-071-838-2 (1-549) x US-08-253-155A-9 (1-8201)
 QY 1 MetAspValValIgluValAlaGlySerTrpTAlaGlnGluArgGluAspIleIleMet 20
 Db 1697 ATGACATGATGAGAGATGAGATGATGTTTCAGGSCACAGAGCGGAAGACATCTTATG 1756
 QY 21 LysTyrGluLysGlyHisArgAlaGlyLeuProGluAspLysGlyProLysProPheArg 40
 Db 1757 AAGTATGACAAAGGACACCGAGCTGGCTGCCAGAGCAAGGGGCTGAGCCCTGTT--- 1813
 QY 41 SerTyrAsnAsnAsnValAspHisIleuGlyIleValHisGluThrGluLeuProLeu 60
 Db 1814 GGAATCAACAGCAGCATGATGCTTTTGGCATTTTGCATGTAGACGAGCTGCCTCGTG 1873
 QY 61 ThrAlaArgGluAlaIleArgLysIleArgArgLysIleSerArgLysSerLysTrpValAsp 80
 Db 1874 ACTGACGCGGAGCGCAAGAAATTCGGCGGAGATGACAGCAAGCAAGAGTGAATGAA 1933
 QY 81 MetLeuGlyAspTrpGluLysTyrLysSerSerArgLysIleuLysIleAspArgAlaTyrLys 100
 Db 1934 ATGCTGGAGAGATGGAGACATATAGACAGCTAGCAAACTCATATGAGCTGTACAAG 1993
 QY 101 GlyMetProMetAsnIleArgGlyProMetLysSerValIleuLeuAsnIleGlnIleMet 120
 Db 1994 GGAATTCCTCATGAACTCCGGGCGCGGTGTGTGCTGCTCTCTGAAACATTTCAGAAATC 2053
 QY 121 LysLeuLysAsnProGlyArgTyrGlnIleMetLysGluLysGlyLysArgSerSerGlu 140
 Db 2054 AAGTTGAAAAACCCCGAGATACACAGATCATGAAAGAGAGGGGCAAGAGGTCACTTGAA 2113

QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
 Db 2114 CACATCCACACACATGACCTGGAGGTGAGACAGACTCCGGAAACCATGCTCTTTAGG 2173
 QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnIleTyr 180
 Db 2174 GATGATATGGAGGCGCAAGAGCAAGAAATTTCTACATCTCTCGGCTATTCGAGATAT 2233
 QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuTyr 200
 Db 2234 AATCCCGAGGTGGCTACTAGCAGGAGCTGAGCCACATACCCGCTGTCTCTCTTAT 2293
 QY 201 LeuProGluLysAspAlaPheTrpAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
 Db 2294 CTGCTGAGAGAGAGACCATTCCTGGGCACTGGGAGCTGTGGCACTGGCACTGAGAGCACTCC 2353
 QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAspGlnGln 240
 Db 2354 CTGCCAGATTCACAGCCCAATGGTGGAGCATCCAGGGGCTCCAAAGCAACAGAG 2413
 QY 241 HisValValAlaThrSerGlnProLysThrMetGlyHisGlnAspLysLysAspLeuCys 260
 Db 2414 CATGTGTATCCCAAGTCAACACCCAAATGGTGGAGCATCCAGGGGCTCCAAAGCAACAGAG 2473
 QY 261 GlyGlnCysSerProLeuGlyCysLeuIleArgIleLeuIleAspGlyTyrSerLeuGly 280
 Db 2474 GGGCAGTGTGGCTCTGTTAAGCTCTCCCGAAGCTGATTTGACGGGATCTCTCTCGGG 2533
 QY 281 LeuThrLeuArgLeuTrpAspValTyrLeuValGlnGlyGlnAlaLeuMetProIle 300
 Db 2534 CTCACTCTGGCTGTGGGCGGTGATTTTGGTGAAGAGCAACAGGTGTGATGCCAATA 2593
 QY 301 ThrArgIleAlaPheLysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyPro 320
 Db 2594 ACCAGATTCCTCTTAAAGTTCAGAGAAAGCGCTCATGAAGACATCAAGATGGCTGG 2653
 QY 321 TrpAlaArgPheCysAsnArgPheValAspThrTrpAlaArgAspGluAspThrValLeu 340
 Db 2654 TGGGACGCTGCGGAAACCAATCTTGATACCTGGGCACTGAGCATGACACCGTCTC 2713
 QY 341 LysHisIleLeuArgAlaSerMetLysLeuThrArgLysGlnGlyAspLeuProPro 360
 Db 2714 AAGCATCTTAAAGGCTCTAGAAAGAAACTTAAAGAAAGCAAGGGGACTGTCACCCCA 2773
 QY 360 ----- 360
 Db 2774 GGCCCAAGCCTTGGAGCGAAGTGTGTGGCAGAAAGCCCCAGCCAGTCTGAACCTTG 2833
 QY 360 ----- 360
 Db 2834 GGGGCACTCCAGAGACCAACCACTGCCCAGCGCTTCCCATGCAAGCAGCAC 2893
 QY 360 ----- 360
 Db 2894 ACCCTTCCTCTGGGATCAGACACTACAGGCTGTGTGCTACAGTCAAGACAGGGCC 2953
 QY 361 ----- 361
 Db 2954 ACACAGAGACCCCAAGACTCCAGAGTCAAGCCCAAAACGGCAAGGGGTCTTGGCAC 3013
 QY 370 rArgProValProAlaSerArgGlyGlyLysThrLeuCysLysGlyAspArgGlnAlaPr 390
 Db 3014 CAGGCTGTGGCGGCTTACGCTGTGGAAAGACCTCTGCAAGGGGTATAGCAAGGCCCC 3073
 QY 390 oProGlyProProAlaArgPheProArgProLysTrpSerAlaSerProProAlaGlnPr 410
 Db 3074 TCCAGGCCCAACAGCCCAAGTCCAGCGGCCCATTTGCTTCCCGCATGGGATC 3133
 QY 410 oArgSerSerThrProCysProGlyValAlaValArgGluAspThrTyrProValGlyThr 430
 Db 3134 TCGTTTTCACGCGCTCTGCTGGGTGGCTGTCCGGGAAACAGATACCTGTGGGCAC 3193
 QY 430 rGlnGlyValProSerProAlaLeuAlaGlnGlyGlyProGlnIleLysTrpArgPheLe 450

DB 3194 TACGGGTGTCGCCAGCTGCTGCTCAGGAGAGACCTCAGGGTTCCGAGATTCTT 3253
QY 450 UGINTTPANSEMERETPROARGLEUPROTHRAPELUAERVALGLUGLYPROTPHAR 470
DB 3254 GAGGTGGAAGTCAATGCCCCGGCTCCCAAGGACCTGGATATGAGGGGCTTGCTCC 3313
QY 470 GHATYAPPEARPEARGLSERCTYETRYVALARGALLESERGLINUAERGLINUA 490
DB 3314 CCAATATGATTTTGAAAGGAGCTGTGCTGCTGCTCATATCCAGAGAGACAGCTGGC 3373
QY 490 APROCYTRGALNALAGLUAHISERPROALAGLU 500
DB 3374 CACCTGCTGCAGGCTGGAACACTGCGGAGAG 3404
RESULT 4
US-09-620-312D-715
Sequence 715, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehtman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghaast
APPLICANT: Drmanac, Radco E.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc FL_genes Version 1.0
SEQ ID NO 715
LENGTH: 1981
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (59)..(1585)
US-09-620-312D-715
Alignment Scores:
Pred. No.: 3,3e-16 Length: 1981
Score: 344.00 Matches: 150
Percent Similarity: 40.18 Conservative: 77
Best Local Similarity: 26.55 Mismatches: 207
Query Match: 11.53 Indels: 131
DB: 4 Gaps: 26
US-10-071-838-2 (1-549) x US-09-620-312D-715 (1-1981)
QY 8 GYSETPTTP-----AAGNGUAATGGLUAERLEIEMETUYTRGLUYS 24
DB 46 GGGCGGCGCTGGATGCGGAGCAAGCAAGAGAAATGGCGCGCGCGCGCGCGCG 105
QY 25 GUYNAGAGLAGLYLEUPROGLUAERLYGGLYPROLYAPROPHARG-SETUYTANYS 44

DB 106 GGAAGCCTGTGCGAAGCCGGAGAGCCTGGCCAGGGCCCGGACCGGCAACCA 165
QY 44 nbnValAerhIsleuGLYLeValhIsleuThrGluLeuProProleuThraAlaArgL 64
DB 166 CGAAGTCAAGCTCTCGG-----TTGACTCGAGAGGCAAGGCTTCCGAGCGCG 219
QY 64 U-----AlaYleGlnI 68
DB 220 CATGCAAGATTGCGCTTCATGCTGGGCTCGAGGGCCCGAGGGCGGCGTGAAGAG 279
QY 68 eaYArgLgLUle-----SerArgYSerLYETRPVALAERMETLEUGLYASPTPGL 86
DB 280 ACCCTGAGCTGTGAGGCAAGAGGAGTCAAGGTGCTGAGCAATGCTTCACAGCTGGA 339
QY 86 uLYETyr-----LYSerSerArgLYLeuLeuLeuAerArgLALYTRLYVGLYMETProMe 104
DB 340 CAATGATATGCGCAAGAACACAAAGAAATGCTGCTGCTGCTCCAAAGAGGATCCGCG 399
QY 104 rAnIleArgGLYPRometTRPSeValLeuLeuAnIleGluMetLYLeuLYAN 124
DB 400 TTCTCTGCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
QY 124 nProGLYArgLYGlnIleMeLYeGLYVGLYVArgSerSerGLYhIsIleGlnAr 144
DB 460 CCTGGAAAGTTTGACGAGCTG---GACATGTCCTGGGAGACCCCAAGTGGCTGAGCT 516
QY 144 gIleAerArgAerValSerGLYThraArgYhIsIleAerPhaArgAArgTyrGL 164
DB 517 GATTAGCTGACCTGCACCGGAGCTCCCATTCATGAAATGTTGTGTCCCGGGGGGG 576
QY 164 YThraYleGlnArgLYLeuLeuHIsIleuLeuAlaYrGLUGLUYUANProGLUYA 184
DB 577 CCAAGGCGGACAGACATTTCCGCTGTAAGGCTTACAGCTGACCGGCGCGGAGGA 636
QY 184 IGLYTYrYrYrAerAerLeuSerhIsIleAlaLeuPhaLeuLeuLYrLeuProGLU 204
DB 637 GGGCTACTGCGGAGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 696
QY 204 uAerAerPhaerThraAlaLeuValGlnLeuLeuAlaSerGLYANhIsIleSerLeuGlnYPh 224
DB 697 GCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
QY 224 hIsIleSerProAerGLYThraValGlnLYLeuGlnIleAerGlnIleValAla 244
DB 748 CTACAGCGAAGAA-----CTGAGGCGATCTCAAGCTGACGCGGAGATCTTTCTC 798
QY 244 aThrSerGln-----ProLYeThrMeGLYhIsleuAerLYVLeuAerLeuCYVGL 261
DB 799 GCTGTTGACAGAGTGTGCGCGGTGCGGCAAGACCTCAGCGCTGAGAAATGCTC 853
QY 261 yGlnCYSerProLeuGLYCYLeuIleArgLleLeuIleAerGLYIleSerLeuGLY 281
DB 854 -----GACCGGCTCTTATATGACAGAAATGTTCAATGCGCTTCTCCGAACTT 906
QY 281 U-----ThraArgLeuThraArgValYrLeuValGlnGLYGLUAlaLe 297
DB 907 GCCCTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 955
QY 297 uMeProIleThraArgLleAlaPhaLeuValGlnIleuYrAerLeuThraYrThSerAr 317
DB 956 -----GTCAGATCATCTTCGCGGTGGCGT----- 982
QY 317 gCYVGLYPROTPRALAARGPhaCYAAnArgPhaValAerPhThraAlaArgAerGLUAs 337
DB 982 ----- 982
QY 337 rThraValLeuLYhIsleuArgLAsErMeLYhIsleuThraArgLYGlnGLYASrPLe 357
DB 983 -GTCTGCTGAAGCAAGCGCTGCGCTCCCTGAAGAGTCAAGCTGCGAGGCGGACAGA 1041
QY 357 uPro-ProProAlaLYrProGLUGlnLYSerSerAlaSerArgProValProAlaSerA 377
DB 1042 CGAAGCATGAGC-----GACTCGGAGGCTTCAAGGCGGCAAGATCA----- 1082

```
QY 377 rggllygllysthrleucylsglyasparaglnalaprobroglyproproalaargp 397
Db 1083 -----TGCAGG----- 1088
QY 397 heProArgProIleTSPser-----AlaSerPro-ProArgAlaProArg--- 411
Db 1089 -----AGGCTTTCGTGTCAGAGAGTGTGTGTCCTGACAGAGCCGAGATTGA 1143
QY 412 SerSerThrProCyseProgllygllyalavalarglualapThrTyrProvalgllythrGln 431
Db 1144 GCGGCAACCTTCATTACGTGCGGCGCTGGCAGAGGACCCG-----GGGTGAGCT 1194
QY 432 GlyAlaProSerProAlaLeuAlaGlnGlnGlyProGlnGlySerTyrArgPheLeuGln 451
Db 1195 GCAGTGCGGCTCCCGCCAGGCT---GCATGTGTCAGAGGCTTCTTGATGCAGAAC 1251
QY 452 TTPAnSerMetProArgLeuProThrAspLeuAspValGlnGlyProTyrPheArgHis 471
Db 1252 TGGTCT---CCGGCTGCTTACAACTTCACACATCCGCTGCGCCCTAGG--- 1302
QY 472 TyrAspPheArgGlnSerCys-----TyrValArgAlaIleSerGlnGlnAspGln 488
Db 1303 -----TGCCTCCCTCCCTGCTCCAAAGCCAAAGCCAAAGCCACCA 1344
QY 489 LeuAlaProCysTTPGlnAlaGlnHisProAlaGlnArgValArgSerAlaPheAla1a 508
Db 1345 GCAGGCGCCAGAGGAGCGGGA---ACAGATGAAAGGAGAGGGGAGCTGGAGAAAGCC 1401
QY 509 ProSerThrAspSerAspGlnGlyThrProPheArgAlaArgAspGlnPro----- 526
Db 1402 CCCAGCCCCCAATACAACTGATGTGTGCGCTGCAGAGATGATGTCCCCACAGCA 1461
QY 527 CysAlaPro 529
Db 1462 TGTGCCCC 1470
```

RESULT 5

```
US-09-774-528-336
; Sequence 336, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Dumanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/774,528
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: P1_F1_genes Version 2.0
; SEQ ID NO 336
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(1476)
US-09-774-528-336
```

Alignment Scores:

```
Pred. No.: 9,686-16
Score: 337.00 Length: 1823
Percent Similarity: 36.23% Matches: 157
Best Local Similarity: 25.28% Conservative: 68
Query Match: 11.30% Mismatches: 231
DB: 4 Gaps: 167
Indels: 29
```

US-10-071-838-2 (1-549) x US-09-774-528-336 (1-1823)

```
QY 16 GluAlaPheIleMetLeuTyrGln---LysGlnHisArgAlaGlnLeuProGluAsp--- 33
Db 154 GAGAACCTGTGTCACCTCCGAGCTGCAGATGATCTACAGCTGCTGGGTCCGATCA 213
QY 34 -----LysGlyProLysProPheArgSerTyrAspAsnAsnValAspHisLeuGlyTle 51
Db 214 GAGCTCAGCGGCTCGGCGCCATATGCTC-----CAGCGCAGCCGTATGATTC 261
QY 52 ValHisGlnThrGlnLeuProProLeuThrAlaArgGlnAlaLysGlnIleArgArgGln 71
Db 262 ATTGGGGGACCTCAGCAGAGCCAGGCGGGCCACCCACCTGCAGACCTCATCCGCA 321
QY 72 IleSerArgLysSerLysTTPValAspMetLeuGlyAspTTPGlyLysSerSer 91
Db 322 -----CGGAGATGAAAGTGTGAGATGATCTCGCAGTGGGAGAAACCATGTCCCG 375
QY 92 Arg-----LysLeuIleAspArgAlaTyrLysGlyMetProMetAsnIleArgGlyPro 109
Db 376 CGGTACAAAGAGTAAAGATGATGATGCGGAAAGGATCCCTGCTCCCTGCGGCCCA 435
QY 110 MetTTPSerValLeuLeuAsnIleGlnGlnMetLeuLeuLysAsnProGlyArgTyrGln 129
Db 436 TGTGTGCGCTGTGTGTGTGGGCGCCATGTGTGCAGAAAGAAAGCCCTGTGAC 495
QY 130 IleMetLysGlnLysGlyLysArgSerSerGlnHisIleGlnArgIleAspArgAspVal 149
Db 496 GAGCTGGCAGAG---GCCCTGGAGAACCCAGTGTATGAGAACCTTGCAGAGAGCTG 552
QY 150 SerGlyThrLeuArgLysHisIlePhePheArgAspArgTyrGlyThrLysGlnArgGln 169
Db 553 CACCGTCATATCCCTGTGCACAGATGTTGTGTGCTCAGAGGCGCACCGGAGCAGAGGG 612
QY 170 LeuLeuHisIleLeuLeuAlaTyrGlnGlnTyrAsnProGlnValGlyTyrCysArgArg 189
Db 613 CTCCTCAGGTGTCTCAAGGCTTACCCCTGATTCACCGAGACAGAGCTACTGCGAGGCC 672
QY 190 LeuSerHisIleAlaLeuPheLeuLeuTyrLeuProGlnGlnAspAlaPheTTPAla 209
Db 673 CAGGGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 732
QY 210 LeuValGlnLeuLeuAlaSerGlnArgHisSerLeuGlnGlnPheHisSerProAsnGly 229
Db 733 CTGTGTCAATATGTGT-----GAGGTCTACTCTCTGTGTGTGTGTGTGTGTGTGT 780
QY 230 GlyThrValGlnGlnLeuGlnAspGlnGlnHisValValAlaThrSerGlnProLys 249
Db 781 -----ATGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 831
QY 250 ThrMetGlnHisGlnAspLysLeuAspLeuCysGlnGlnCysSerProLeuGlnCysLeu 269
Db 832 CTGCTTCGACAGTGCACAAAGCAGTGCAGAGTGGCGGTGGAGCCCTGTGTACTG 891
QY 270 -----LysArgIleLeuIleAspGlyIleSerLeuGlnLeuThrLeuArgGln 285
Db 892 CCCGAGTGTCTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 951
QY 286 TTPAspValTyrLeuValGlnGlnGlnAlaLeuMetProIleThrArgIleAlaPhe 305
Db 952 TGGATGCTCTCTCTCAGTGAAGGATGC---CAGAGTACTGTTCG----- 992
QY 306 LysValGlnGlnLysArgLeuThrLysThrSerArgCysGlyProTTPAlaArgPheCys 325
Db 993 -----TGTGG----- 998
```


Db 502 CTGCGCGGACATCCCATGGCATGAGCCACAGACTGTGATCGGCTCTGTGGGCGCTG 561
 QY 119 G1umeTyLeuYsaAnProG1YaTyYr---G1n1LeMeTyLeuYsaG1YLaTyArg 137
 Db 562 CAGAGAGAGAGAACTGTAGCTGTCTCTACCCGAGATGTGTAGAGAAAGCTCCAGCAT 621
 QY 138 SerSerGluHis1LeG1nArg1LeAsPArgAsPValSerG1YThrLeuArgYsaHis1Le 157
 Db 622 GAGACCATCGCTGCCAAGCATCGAAGAGAGCTGTCTCGGACCATGATGCCCAAGACGCC 681
 QY 158 PhePheArgAsPArgTyC1YThrYsaG1nArgYsaLeuHis1LeLeuLeuAlaTyYr 177
 Db 682 TGCTTCGCGAGATGAGTATGATCGGAGTGCCTGTGCGAGGAGGCTCGGCGCTG 741
 QY 178 G1uG1uTyYsaAnProG1uValG1YTyYrCyAsArgAsPLeuSerHis1LeAlaLeuPhe 197
 Db 742 GCTGCTCTTACCCAGAGATCGGCTACTGCCAGGAGCAGCGGATGTGTGCGCTGCTGCTG 801
 QY 198 LeuLeuTyYrLeuProG1uArgYsaPheTyP-----AlaLeuValG1n---Leu 213
 Db 802 CTGCTGTCTCTGAGAGAGAGAGAGAGCTGTCTGTGATGATGTGTGCATATCGAGAGACCTG 861
 QY 214 LeuAlaSerG1uArgYsaHisSerLeuG1nG1YPheHisSerProAsnG1YThrValG1n 233
 Db 862 CTCCCGGCTCTACTTCAGACAC-----G1uHisValValAlaThrSerG1nProYsaThr 250
 QY 234 G1YLeuG1nAsPArg1nG1n-----G1uHisValValAlaThrSerG1nProYsaThr 250
 Db 895 GGATGTCAGACTACACAGCGGCTCTGCGCACCTCATGTGTCCAGTACCTGCTGCGCTG 954
 QY 251 MetG1YHisG1nAsPArgYsaPLeuCyAsG1nG1YnCySerProLeuG1YCyYsaLeu1Le 270
 Db 955 GACAGCTCTCTCGAGAGATGATGATGAGCTGTCTGATGATGATGATGATGATGATGATGAT 1014
 QY 271 Arg1Leu1Leu1AsPArg1Y1LeSerLeuG1YLeuThrLeuArgYsaPValTyYrLeu 290
 Db 1015 ACAGGCTTGTGCGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074
 QY 291 ValG1uG1YsaG1nAlaLeuMeTyPLeuThrArg1LeuAlaPheYsaValG1nG1nYs 310
 Db 1075 TACAGAGGCTCTCGGAGCTGTCTTCACGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 1134
 QY 311 ArgLeuThrYsaHisSerArgYsaG1YProThrAlaArgPheCyAsAsnArgPheValAsP 330
 Db 1135 GAGCTGATTCAGCTCAGAGAACTCG-----GCCTTCATCTTCAACACGCTATCGGAT 1185
 QY 331 ThrThrAlaArgAsPArgYsaPThr-----ValLeuYsaHis1LeuArgYsa 345
 Db 1186 ATCCCGTCGAGATGAGAGCGGAGCTGTCTTGGGGGTGGCCATGCGGCTGCGCGC 1245
 QY 346 SerMeTyLeuYsaLeuThrArgYsaG1YsaPLeuProP----- 359
 Db 1246 TCCTCTAC--CGATGTGCGCT-----GAGACTCAGCGCGCAAGCTCTGCGCTATCT 1298
 QY 360 -----ProAlaYsaProG1uG1nG1Ys 367
 Db 1299 CATTGACAGACAGGCGAGCTCTTGGGGCGGACCTTCAACCACTCTCTCAGGTTGT 1358
 QY 368 SerAlaSerArgProVal-----ProAlaSerArgYsaG1Y--- 379
 Db 1359 TCGCGCAGAGACCGAGAGAGTCCACATCATCTCTCTCTTGGGGAGAGATGA 1418
 QY 380 -----LysThrLeuCyAsYsaG1YAsPArgYsaValProProG1Y 392
 Db 1419 CTTGAGGACACTCAAGCGCAAGAACATCAAGACAGAGCTGTGTGCTGCTGCGGGA 1478
 QY 393 Pro---ProAlaArg---PheProArgPro1LeuThrPheHisSerProProArgYsaPro 410
 Db 1479 AGCCATCTGTGCGCTGCGACGCTTCCAGT-----CACAGACCCCAAAACTGCGAG 1532
 QY 411 ArgSerSerThrProCyAsPProG1YsaValaValaArgYsaPThrTyProValG1YThr 430
 Db 1533 CGTGTGATGTGCGACGCTCTGCGCTGCT-----ACC 1565

QY 431 G1nG1YsaPProSerProAlaLeuAlaG1nG1YsaPProG1n-G1YSerThrPAsPheLe 450
 Db 1566 AAACAGCGGCTTACTTACCTTCAACCCCTTGGTGGGCTGTGTTCATCTGTGCGAGAGCT 1625
 QY 450 uG1nThrAsn---SerMetProArgYsaPProThrAsP----- 461
 Db 1626 GACTCCAGACTATAGATATGAGAGACCAAGCGGAGCAGAGAACTACGAGGCTGCTC 1685
 QY 462 -----LeuAsPValG1uG1YProThrPheArgYsa 471
 Db 1686 ACCGACCAACGCGCGCCGAGCCAGCCCTCTGACTTTGAG-----CGGCA 1733
 QY 471 sTyYrAsP-----PheArgYsaSerCySTPValArgAla1LeSerG1nYsa 487
 Db 1734 CGACGACGACGAGCTGCGCTTCCGAGAGAGACATCATCATCATGCTGTCTCAGAGAG 1793
 QY 487 pG1nLeuAlaProCySTPTrG1nAlaG1u-----HisProAlaG1 500
 Db 1794 CGAG-----CACTGTGCGGTGGGAGCTCAACGCGCTGCGAGAACTAGCTGCTGCT 1847
 QY 500 uArgYsaArgSer---AlaPheAlaPProSerThrAsPserAsPArgYsaPThrPro 519
 Db 1848 GTTGTGGAAGTCTGTGATGAGCGCAGCAAGAGTACTTCATCTGCGGGGATGACTCGT 1907
 QY 519 heArg-----AlaArgAsPArgYsaPLeu 535
 Db 1908 GACCGAGGGGCTCAGACCTGTGTGAGAGACCTCTGCGGCGCTTAAAGGCTGT 1967
 QY 528 laPro-----ThrSerG1YProCyYsa 535
 Db 1968 CGAACAATGACATGAGAAGCATCCCTGCTT 1999

RESULT 7
 US-09-620-312D-676
 / Sequence 676, Application US/09620312D
 / Patent No. 6569662
 / GENERAL INFORMATION:
 / APPLICANT: Tang, Y. Tom
 / APPLICANT: Liu, Chenghua
 / APPLICANT: Asundi, Vinod
 / APPLICANT: Zhang, Jie
 / APPLICANT: Ren, Feiyan
 / APPLICANT: Chen, Rui-hong
 / APPLICANT: Zhao, Qing A.
 / APPLICANT: Wehrman, Tom
 / APPLICANT: Xue, Aidong J.
 / APPLICANT: Yang, Yonghong
 / APPLICANT: Wang, Jiao-Rui
 / APPLICANT: Zhou, Ping
 / APPLICANT: Ma, Yundong
 / APPLICANT: Wang, Dunrui
 / APPLICANT: Wang, Zhiwei
 / APPLICANT: John Tillinghast
 / APPLICANT: Drmanac, Radoje T.
 / TITLE OR INVENTION: No. 6569662el Nucleic Acids and
 / TITLE OF INVENTION: Polypeptides
 / FILE REFERENCE: 784CIP2B
 / CURRENT APPLICATION NUMBER: US/09/620,312D
 / PRIOR APPLICATION DATE: 2000-07-19
 / PRIOR APPLICATION NUMBER: 09/552,317
 / PRIOR FILING DATE: 2000-04-25
 / PRIOR APPLICATION NUMBER: 09/488,725
 / NUMBER OF SEQ ID NOS: 1105
 / SOFTWARE: pL_FL_genes Version 1.0
 / SEQ ID NO 676
 / LENGTH: 2955
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (169)..(2418)


```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (1) (2955)
OTHER INFORMATION: n = a,c,t,c or g
US-09-620-312D-676

```

Alignment Scores:

Pred. No.:	3, 06e-09	Length:	2955
Score:	253.00	Matches:	156
Percent Similarity:	35.80%	Conservative:	71
Best Local Similarity:	24.61%	Mismatches:	233
Query Match:	8.48%	Indels:	175
		Gaps:	28

US-10-071-838-2 (1-549) x US-09-620-312D-676 (1-2955)

```

QY 9 SerTPTRPAlaGlnGluArgGluAerPLeileMetLeuYr---GluYvGlyHisArg 27
DB 220 ACCATATGGCCCGCAGAG-ATCTGGCCAGTACACGAGAAAGAGATCA 270
QY 28 AlaGlyLeuProGlu-----AerYvGly 35
DB 271 GAGAGCAACAGAGTTCTACTACGATGATTGTTCCGTGTACAGAGAAAGAT 330
QY 36 ProYvProPheArgSerTyraAerAerAerAerAerAerAerAerAerAer 47
DB 331 GATGACCTGGCTCCGCTGCTGCGGAACTCCCTGTATGAGAGATGCTCCACAGAG 390
QY 48 -----HisLeuGlyLeuValHisGluThrGluLeuProLeuThrAla 62
DB 391 CTGGGTGAGAGCCACCTGAGATTCACTCATACAGATGAGGAGATCTCAC-- 447
QY 63 ArgGluAlaLeuGlnLeuArgGluLeuSerArgYvSerTyraValAerMetLeu 82
DB 447 ----- 447
QY 83 GlyAerTPGluYv-----TyraYvSerAerArgYvLeuLeuAerAla 98
DB 448 -----TGGACAAAGATTCGCTCCCTCAACCCGCTCGAAGAGCTCCGCTCCGCTG 501
QY 99 TyraYvGlyMetProMetAerAerAerAerAerAerAerAerAerAerAerAer 118
DB 502 CTGGCGGATCCACATGATGAGAGCCACAGCTGTGATGCGGCTCTGCGGCGCTG 561
QY 119 GluMetLeuLeuAerProGlyArgTyra---GlnLeuMetLeuYvGlyYvArg 137
DB 562 CAGAAAGAGAGAACTGATGCTCTCCGCGAGATTTGAAAGAGCTCCAGAT 621
QY 138 SerSerGlnHisLeuGlnAerAerAerAerAerAerAerAerAerAerAerAer 157
DB 622 GAGACCATGCTGCGCAAGATCGAAGAGAGCTGCTCGCACTGCCAGCAAGCC 681
QY 158 PhePheArgAerArgTyraGlyThrYvGlnArgGluLeuHisLeuLeuAlaTy 177
DB 682 TCGTTGCGAGATGAGATCGGAGTGGCCGCTGCGGAGGCTCGCGGCGCTG 741
QY 178 GluGluTyraAerProGluValGlyTyraArgAerLeuSerHisLeuAlaLeuPhe 197
DB 742 GCGTGGCTTACCCAGAGATCGGCTTCTGCGAGGAGCCGCGATGCTGCGCTG 801
QY 198 LeuLeuTyraLeuProGluGluAerAerAerAerAerAerAerAerAerAer 213
DB 802 CTGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 861
QY 214 LeuAlaSerGlnAerHisLeuGlnGlyPheHisSerProAerGlyLeuThrValGln 233
DB 862 CTCCCGGCTCCCTTCAAGACCC-----ACCTGCTG 894
QY 234 GlyLeuGlnAerGln-----GlnHisValValAlaHisSerGlnProYvThr 250
DB 895 GGTGTCACATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
QY 251 MetGlnHisGlnAerYvAerLeuYvGlyGlnCysSerProLeuGlyCysLeuIle 270

```

```

DB 955 GACAAGCTCTCCAGAGATGACATTGAGCTGCTCCATGATCACTGACATGTTCTC 1014
QY 271 ArgGlnLeuLeuAerPLeileSerLeuGlyLeuThrLeuAerGluAerPLeile 290
DB 1015 ACCGCTTTCGCGAGGCTGTGGATCAAGATGCTCTCTGCGATGCGGCTGTTTC 1074
QY 291 ValGlnGlyGlnGlnAlaLeuMetProLeuThrArgLeuAlaPheLeuValGlnGly 310
DB 1075 TACAGAGGCTCCCGGTGTGTTCACAGCTGACCTGCGGATCTGCACTTCAAGAGAA 1134
QY 311 ArgLeuThrYvThrSerArgCysGlyProTPRAlaArgPheCysAerAerPheValAer 330
DB 1135 GAGCTGATCCAGTCAAGAACTCG-----GCTCATCTTCAACAGCTATGAGAT 1185
QY 331 ThrTPRAlaArgAerGluAerThr-----ValLeuLeuHisLeuArgAla 345
DB 1186 ATCCCGTCCAGATGAGAGCCGAGAGCTGTTGAGGAGTGCATGCGGCTGCGCGG 1245
QY 346 SerMetLeuYvLeuThrArgYvGlnGlyAerLeuProProAerAerProGluGln 365
DB 1246 TCCCTCAC-CGATGCGCT-----GAGACTCAGCGCCGAGACT-----CCT 1289
QY 366 GlySerSerAlaSerArgProValProAlaSerArgGlyYvLeuThrLeuCysYvGly 385
DB 1290 GGCCTATCTCATGACAGACAGAGCGAGCT----- 1319
QY 386 AerPArgGlnAlaProProGlyProProAlaAerPheProAerProLeuThrSerAlaSer 405
DB 1320 -----CTGAGGCGCGAGCCCTTCAACACCTTCTGAGGATGT 1358
QY 406 ProProArgAlaProArgSerSerThrProCysProGlyValAlaValArgGluAerThr 425
DB 1359 TCGCCGAGAGACCCAGCGAGAGATCCACATCACTGCTGCTCTTGG--GAGAGA 1415
QY 426 TyrProValGlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyYv---ProGln 444
DB 1416 TGACCTGAGAGACATCAAGAGCCAAAGATCAAGAGAGAGAGAGAGAGAGAGAGAG 1475
QY 445 GlySerTP-----ArgPheLeuGlnTP 452
DB 1476 GGAAGCGATCTCGCGTGGAGCCGATCTTCAAGTGCACAGACCCCAAACTGCGAGCT 1535
QY 453 AerSerMetProArgLeu-----ProThr-Aer----- 461
DB 1536 GAGAGCTACTCAAGATTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1595
QY 462 -----LeuAerValGlnGlyProTPR 469
DB 1596 GTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1644
QY 469 AArgHisTyraP-----PheArgGlnSerCysTPRValArgAlaLeuSerGln 485
DB 1645 -CGGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703
QY 485 nGluAerGlnLeuAlaProCysTPRAlaGln-----HisP 498
DB 1704 GAGAGAGAG-----CACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1757
QY 498 AlaGluArgVal-ArgSer--AlaPheAlaProSerThrAerSerAerGlnGly 517
DB 1758 AGCAAGATTCGAGAGATCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1817
QY 517 hTPRProPheArg-----AlaArgAerGluGlnP 526
DB 1818 CTGCGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1877
QY 526 roCysAlaPro-----ThrSerGlyProCysLeu 535
DB 1878 CCGTTCGAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1915

```

RESULT 8

US-09-248-796A-5457


```

QY 157 llepPheArg-----AspArgTyrGlyThrLeuGlnArgGluLeuHisIleLeu 174
DB 504 CCGGCAATTTCAGAACCCGATGCGATAGGTCCTCGCGAGA-----GTCCTT 551
QY 175 LeuAlaTyrGluGlnIleThrAnpProGluValGlyTyrCysArgAspLeuSerHisIleAla 194
DB 552 CAGGCTTATCGCTGCGGATCCAGATCGAGTGGCTACTGTCAGGCCATGAAATATTGTGTG 611
QY 195 AlaLeuPheLeuLeuTyrLeuPheProGluGlnAspAlaPheThrAlaLeuValGlnLeuLeu 214
DB 612 TCGGTGTTCTCTCTCTCTGTGATGAGAGAACGATTCGTGATGTCGCCAGCCCTGTGTC 671
QY 215 AlaSerGluArgHisSerLeuGlnGlyPheHisSerProAsnGlyIleThrValGlnGly 234
DB 672 -----GAGAACTTACTTCCGCACTTACACAGATAAA-----GTGGGGGC 713
QY 235 LeuGlnAspGlnGln-----GluHisValValAlaThrSerGlnProIleThrMet 251
DB 714 GCCCAATCGATCAGGGGTGTGCTTAATGAGCTAGTGAGACGCAATTGCTGATTGGCAC 773
QY 252 GlyHisGlnAspIleValAspLeuLeuCysGlyGlnCysSerProLeuGlyCysLeuIleArg 271
DB 774 GGCCATCTGAGCAGATGGGTGTGATTAAGATGATTTC---ATTTCCTGTTCTCTTACC 830
QY 272 IleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuThrAspValTyrLeuVal 291
DB 831 ATCTTCATGAGGTGATTCAGCTACGAGAGCTGCTGACATCTGAGCTTTTTC 890
QY 292 GlnGlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheIleValGlnGlnIleArg 311
DB 891 GAGGGGCGCAGATCATTTTATGATTTCACCTGCAGATCATTAATGAGAACGGGACAA 950
QY 312 LeuThrIleThrSerArgCysGlyProThrAlaArg-PheCysAsnArgPheValAsp 331
DB 951 CTGTGATCTGCCAAGATGATGCGCAAGCCATGCTAGTTTGCAGA-----ACTAC 1001
QY 331 rTrpAlaArgAspGluAspThrValLeu----- 340
DB 1002 CTGGAGG-----GAGTTTCAATCCAGAGTACAGAGTGCCTCTTACGAGCAAAACGG 1055
QY 341 -----LysHisLeuArgAlaSerMetIleValLeuThr 352
DB 1056 AAGATGAGCGTAAAGTCCAAACAGAAACAGTTCAACCTGATTCATGAGGATACACA 1115
QY 352 GlySerGlnIleAspLeuProProAlaIleValAspProGlnGlnIleSerSerAlaSerArg 372
DB 1116 AAAAGATTCGAT-----GAGGAGTATGACACAGCAGCGCATCGAGAG 1160
QY 372 cValProAlaSerArgGly----- 379
DB 1161 TTAACGACACAGATGCGCGTTAACGATGCGACAGTTCGATATGCAACAGAAAGACC 1220
QY 380 -----Lys 381
DB 1221 ATTGTCAAGGCTTACGTCAGATCGTATTTCACCGCAGTGAATGTCATATGTCGTC 1280
QY 381 rLeuCybIleValAspArgGlnAlaProProGlyPro----- 393
DB 1281 ACTATCATCTCGGAGAGAAAGCAGCGCTTAATCTTCCTGCAACAGCAGCAGCAAGATT 1340
QY 394 -ProAlaArgPheProArg---ProIleThrSerAlaSerProProArgAlaProArg 412
DB 1341 CAATGCGCGTGTCTGAGACTCCCGCAGCTGTCCTCACT---CGACAAAGCAGACCGATC 1397
QY 412 rSerThrProCybProGlyIleValAlaValIleAspThrTyrProValGlyThrGlnGln 432
DB 1398 CAAGATGCGCGGCGCTCC--GAGAGC---AGATACGAGGCTTACAGTCAAGTCAATGA 1452
QY 432 yValPro-----SerProAlaLeuAlaGlnGlnGlyIleProGlnIleSer 447
DB 1453 TATTCACACACTTTTACGAGAACTGACACCCCTGCTAAGTCAAGTCAAGTCAAGTCA 1512
QY 447 pArgPheLeu-----GlnThrAsnSerMetProArgLeuProThrAs 461

```

```

DB 1513 GCGAAATAATGTTTAGCTTACGATTAAGAGAGAACTGTGTCTTCGACTTCGACAGC 1572
QY 461 pleuAspValGlnGlyPro 467
DB 1573 TCATCAATGATATGGGCTT 1591

RESULT 10
US-09-270-767-9899/C
; Sequence 9899, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7325-094
; CURRENT FILING DATE: 1999-03-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9899
; LENGTH: 806
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9899

Alignment Scores:
Pred. No.: 1,056-05 Length: 806
Score: 194.50 Matches: 79
Percent Similarity: 41.14% Conservative: 44
Best Local Similarity: 26.42% Mismatches: 111
Query Match: 6.52% Indels: 67
DB: Gaps: 9

US-10-071-838-2 (1-549) x US-09-270-767-9899 (1-806)
QY 101 GlyMetProMetAsnIleArgIleProMetTyrPheValLeuLeuAsnIleGlnIleMet 120
DB 788 GCGGTTCCAGAAAGCACTGCGTGAAGAAAGATCTGCGAGAACTGACCAACGTCGAGGCG 732
QY 121 LysLeuLysAsnProGlyIleArgTyrGlnIleMetLysGlnGlyIleValArgSerSerGlu 140
DB 731 AGAATGAATGAACAGCAGCAAGTAAAGATCTTAATCCAAAGAAACCAAAAGTGAAGCC 672
QY 141 HisIleGlnArgIleAspArgAspValSerGlyThrLeuArgLysHisIlePhePheArg 160
DB 671 GTTATTCAG-----CGGACATCCATCGCACTTTCGCGGCGCAAAATGCTTCAAA 621
QY 161 AspArgTyrGlyThrLysGlnArgGluLeuLeuHisIleLeuLeuAlaTyrGlnGlnIle 180
DB 620 GAAATTTGGCGGTTCTGCGCCAAAGATGCACTTTTAAGTGTCCAAAGCGTATGCCCTTCA 561
QY 181 AsnProGluValGlyTyrCysArgAspLeuSerHisIleAlaAlaLeuPheLeuLeuTyr 200
DB 560 GACAGCAGAGTGTGATATGTCAGGCTTAATTCATATGACAGCTGTCTGCTTCAAT 501
QY 201 LeuProGlnGlnAspAlaPheThrAlaLeuValGlnLeuLeuAlaSerGluArgHisSer 220
DB 500 ATGCTCAGAGAGATGCTGTCTGTCTGTGTAAGCGCTTATGATGACAC----- 453
QY 221 LeuGlnGlyPheHisSerProAsnGlyIleThrValGlnGlyLeuGlnAsp----- 237
DB 452 -----TATGGGCTCGGTATCTACAAA 429
QY 238 -----GlnGlnGluHisValAlaThrSer 246
DB 428 GCCGATTCGAGATCTTACCTGCTCTCTCTCAACATCGAGCGATGCAAGATCAG 369
QY 247 GlnProLysThrMetGlyHis-GlnAspIleValAspLeuCysGlyGlnCysSerPro 266
DB 368 CTGCGCAAGCTGCAAGAACACTTTCACGCGCTCGGCTTGAAGACGACATGATGCGCTCC 309
QY 266 uGlyCysLeuIleArgIleLeuIleAspGlyIleSerLeuGlyLeuThrLeuArgLeuThr 286

```

Db 308 CAATGG-TTCCTAACCCTATATACAGCGCATCCATTGTTGTTGTTCCAGTGCT 250
Qy 286 PAAPVALTYRLEUVALGLUGLNALALEUWETPROILETHRARGLEALAPHELY 306
Db 249 GGAAGTGTCTTACTGACGGA-----CTACCTGTG----- 219
Qy 306 svaIaGlnGlnLYARGLEUThrLYRThrSerARgCySGLYPROTPRALARGPhECyAAs 326
Db 218 -CTCTCCAGGTGGCGGTTACTCTCTTATCATCTGTGATCC----- 177
Qy 326 nArgPheValaSPThrTTPRALARGAspGlnuSPThrValLEuLYSHILEUARGAlase 346
Db 176 -----GACTTGGCGGCAACTCGATTGTGAGGCGCATTTAAAGTATTTCCGGGTAC 127
Qy 346 rMetLYsLYsLeuThrARGLYSGInGLYAsPLeuPROPROAlaLYSPROGLUGInGL 366
Db 126 GTTGGCGGAAAAGTGGCGGAGCTCC-----ACTCAGGC-ACGCAAG 86
Qy 366 ySerSerAlaSer-----ArgProValPROAlaSerARGLYGLY 379
Db 85 TGATGAAGCAAGCTCTGCGAACGTAAAGATAAAAAAGTGAAGCAATGACGAGAGCA 31

RESULT 11
US-08-363-300-1
Sequence 1, Application US/08363300
Patent No. 5700927
GENERAL INFORMATION:
APPLICANT: Zon, Leonard and Richardson, Paul
TITLE OF INVENTION: Tpol Gene and Uses Thereof
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,300
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 04590/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4039 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 403..3829
US-08-363-300-1

Alignment Scores:
Pred. No.: 0.000768
Score: 182.50
Percent Similarity: 34.59%
Best Local Similarity: 22.52%
Query Match: 6.12%
DB: 1
Length: 4039
Matches: 125
Conservative: 67
Mismatch: 227
Indels: 139
Gaps: 18

US-10-071-838-2 (1-549) x US-08-363-300-1 (1-4039)
Qy 13 GlnGlnuArgGlnuAspIleIle-----MetLYsTYRGLULYsGLYHISARGAlaGLY 29
Db 2408 AGAAAAGAGAAAGACCTCAGCAGGCTTCAGAGGCTGTGAAAAGGCCATTTCCAGCA 2467
Qy 30 LeuPROGlnuSPryGLY---ProLYSPROPhelARGSerTYRAsnAsnValaSPHIS 48
Db 2468 GATCTCGCTCCGTCAGAGATGAGAGAGATCAACAACCTCAAG-----CCT 2515
Qy 49 LeuGLYIleValHISGLInThrGLULeUPROPROLeuThrAlaARGuAlaLYsGlnIle 68
Db 2516 CTGAAAACCATTTGCTGAACCAACCTCAAGCTTGAATAGAAATCACTCCGTTC 2575
Qy 69 ARGARGInuIleSerARGLYsSerLYsTYRValaSPMetLeuGLY----- 83
Db 2576 TT-AAAGAAATCAGT-----ACAGTGGGAAAAGATGCTTACCACTCCAGAAAGATCC 2628
Qy 84 -----AsPTRGInuLYsTYRLYsSerSerARGLYsLeuIleAspARGAlaTYR 99
Db 2629 AAAATTAAGTTTGAATGGAAGAAAAGTGACACTCAGCT-----GTTGGG 2670
Qy 100 LYsGLYMetPROMetAsnIleARGLYPROMetTYRSPeValLEuLeuAsnIleGLInu 119
Db 2671 CAGGTGTGCCACGTCATCAGCAGGAGTGAGATCTGAAATTTCTAGCTGAGCAGATTCCAC 2730
Qy 120 MetLYsLeuLYsAsnPROGLYARGTYRGLN-----IleMetLYsGLYULYsGLY 135
Db 2731 CTAAACACCATTTCTCAGTAAGAAACAGACCAAGACAGCTGCCCTCAAAAGAGCTCTTG 2790
Qy 136 LYsARGSerSerGLInuHISIEGLInuArgIleAspARGsPValaSerGLYThrLEuARGLYs 155
Db 2791 AAGAGCTGACCTTGGCAGCAGCAGCGCCATTCATCGACCTCGGGGAAACCTTTCCACA 2850
Qy 156 HISIlePhePheARGAspARGTYRGLYThrLYsGlnuArgGlnuLeuHISIEleuLeu 175
Db 2851 CATCCATACCTTCTGTGCCAGCTTGGAGCAGGCTCAGCTGCTATTAACAACATTCGAAG 2910
Qy 176 AlaTYRGLInuLYsAsnPROGLInuValGLYTYRQYsARGAsPLeuSerHISIElaIa 195
Db 2911 GCTTACTCGCTTGTGACCAAGAGGTTGATCTGCAAGCTGCTGAGCTTTGTGGAGGC 2970
Qy 196 LeuPheLeuLeuTYRLeuPROGLInuAspAlaPheTPRALeUValGlnLeuAla 215
Db 2971 ATTGTGCTTTCATCATGAGTGAAGAGGGGTTTAAAGTCTCAAGTTCCTGATGTTT 3030
Qy 216 SerGLuARGHISerLeuGlnLYsPheHISerPROAsnGLYGLYThrValGlnGLYLeu 235
Db 3031 GAC-----ATGGGGCTGGGAAACAGATCGGCCAGACATGATTAATTTCAGATCCAG 3084
Qy 236 GlnAspGLInGLInuHISValaValaIaThrSerGlnPROLYsThrMetGLYHISGlnAsp 255
Db 3085 ATGTACCACTGTGACAGGCTCTTCCACAGATTACCAAGACCTTCAACACCTGGAA 3144
Qy 256 LYsLYsAsPLeuCYsGLYGLInCYsSerPROLeuGLYCYsLeuIleArgIleLeuIleAsp 275
Db 3145 GAGCAGAGACTGGCCCCCTACGTACGCGGCTCCGTGTTTCCACCGTTCGCCCTCA 3204
Qy 276 GLYIleSerLeuGLYLeuThrLEuARGLeuThrAspValTYRLeuValGlnGLYGLIn 295
Db 3205 CAGTTCACCTGCGCTTTTGAAGCCAGAGTCTTTGATATAGATCTTCTTCAAGGATCAGAG 3264
Qy 296 AlaLeuMetPROIleThrARGIleAlaPheLYsValGlnGLInLYsARGLeuThrLYsThr 315
Db 3265 -----GTCAATTTTAAAGTACCTTAAGCTTTTG----- 3294
Qy 316 SerARGCYsGLYPROTPRALARGPhECyAsnARGPheValaSPThrTPRALARGAsp 335
Db 3295 -----GGAGGCAAT 3303
Qy 336 GlnAspThrValLEuLYsHIS-----LeuARGAlaSerMetLYsLYsLeuThrARGLYs 353
Db 3304 AAGCCCTTGATTTTACAGCATGAGAACTGGAACCAATCGTGAATTCATTAAGAACACA 3363

```

QY 354 GINGLYAspLeuProProAlaLysProGluGlnGlySerSerAlaSerArgProVal 373
DB 3364 CTCGCCAACCTGG-CCTGGTCAGATGAGAAAGACATCACTCAAGT----- 3410
QY 374 ProAlaSerArgGlyGlyLysThrLeuCyLysGlyAspArgGlnAlaProProGlyPro 393
DB 3411 -----GTTTGAGATGACATCGCCAAAGACACTCAGAGCCCTTA 3446
QY 394 ProAlaArgPheProArgProIleThrSerAlaSerProProArgAlaProArgSerSer 413
DB 3447 TGAGGTGCGAGTACCAAGTCTGT----- 3467
QY 414 ThrProCyPProGlyGlyValAlaArgGluAspThrThrProVal----- 428
DB 3468 -----CCAGAGAGAGACT-----TATTGAGTCTCTCCCTCTCACTGA 3503
QY 429 -----GlyThrGlnGlyValProSerProAlaLeuAlaGlnGlyProGln 444
DB 3504 CAACCAAGAAATGAGAAATGAGAAACCAACAGCAGCTTGGCAACAGAACTTG- 3562
QY 445 GlySerTPArgPheLeuGlnTP----- 452
DB 3563 ACCTCTGGAGCAGTTGCAAGTGCATAATGCTAGATCCAAAGCCTTGAAGCCAGCTAG 3622
QY 453 ---AanSerMetProArgLeuProThrArgPheLysValGlnGlyProThrPheArgHis 471
DB 3623 AGAACTTCTTACCAAGCAGAGTACGCTGACAGCAGCTGCTGACCTGG----- 3673
QY 472 TyrAspPheArgGlnSerCySerTPValAlaArgAlaIleSerGlnGlnLysPheLysAlaPro 491
DB 3674 -----AGGTGAGAGCGTGGCTGCTGCTGAG----- 3697
QY 492 CySerTPGlnAlaGlnIleProAlaGluArgValArgSerAlaPheAlaAlaProSerThr 511
DB 3698 -----ATGTTGGAGAGAGCTGCGAGGCAAAAGCCCGCCAGCAGACT 3739
QY 512 AspSerAspGlnGlyThrProPheArgAlaArgAspGlnPro 526
DB 3740 CCAGAGCCAGACTGCACCACTGAGCCCAAGCCGATTGACCG 3784

RESULT 12
US-09-248-796A-1133
; Sequence 1133, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1133
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1133

Alignment Scores:
Pred. No.: 0.000707 Length: 1056
Score: 172.00 Matches: 49
Percent Similarity: 47.18% Conservative: 43
Best Local Similarity: 25.13% Mismatches: 67
Query Match: 5.77% Indels: 36
DB: 4 Gaps: 7

US-10-071-838-2 (1-549) x US-09-248-796A-1133 (1-1056)

```

```

QY 145 IleAspArgAspValSerGlyThrLeuArgGlyHisIlePhePheArg----- 160
DB 28 ATTGAAGAGATTGATGATGATCTTCCCTGATATATATCTTTAATTCATCAATCAAA 87
QY 161 -----AspArgTyGlyThrLysGlnArgGlu----- 169
DB 88 GGTGTGTTTTTCATCTTGAAGAACTTGGCAACAAAGAAAGAAACATTAATGATC 147
QY 170 -----LeuLeuHisIleLeuLeuAlaTyGluGluLysTyAsnProGlnValGlyTyCyS 187
DB 148 AAGCATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
QY 188 ArgAspLeuSerHisIleAlaLeuPheLeuLeuTyLeuProGluGlnAspAlaPhe 207
DB 208 CATGATTAATTAATTTTGGCAGCTTTTATTTATTTGTTTCATGGAAGAAAGAAACATTT 267
QY 208 TPAlaLeuValGlnLeuLeuAlaSerGluArg-----HisSer 220
DB 268 TGGATGCTTGT-----ATATTAACAGAAAGAAATTAATCCAAAGTTTCATCTGCAAT 321
QY 221 LeuGlnGlyPheHisSerProAsnGlyGlyThrValGlnGlyLeuGlnAsp----- 237
DB 322 CTGAGAGGTGTTCAATCACTGATCAAGGCTTTAAATGTTATGTTTAAAGATATATACC 381
QY 238 GINGLInGlnIleValAlaIleThrSerGlnProLysThrMetGlyHisGlnAspLysLys 257
DB 382 CAATTATGCAAGTTTACGAGAAATTTCCATGATGAAACCTTTCTGAAACAAAG---- 438
QY 258 AspLeuCyGlyGlnCySerProLeu-----GlyCyLeuIleArgIle 272
DB 439 ---ATTATTAACAAGATTACCTTCAGTTACATTAGTCACTTCATTGTTGATGCAGTA 495
QY 273 LeuIleAspGlyIleSerLeuGlnLysLeuArgLeuTPAspValTyLeuValGlu 292
DB 496 TTGTGGGATTTCTACTTATGAAACCACTTAAAGATTTAGGATATCTTTGGTATGAA 555
QY 293 GlyGlnGlnAlaLeuMetProIleThrArgIleAlaPheLysVal 307
DB 556 GGTGGAACAATTTTCAGATTCTCTTAACCATATTCAAATG 600

RESULT 13
US-09-248-796A-1820
; Sequence 1820, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1820
; LENGTH: 1659
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1820

Alignment Scores:
Pred. No.: 0.0284 Length: 1659
Score: 154.00 Matches: 46
Percent Similarity: 44.84% Conservative: 54
Best Local Similarity: 20.63% Mismatches: 111
Query Match: 5.16% Indels: 12
DB: 4 Gaps: 4

US-10-071-838-2 (1-549) x US-09-248-796A-1820 (1-1659)

```

QY 78 TPValAspMetLeuGlyAspTPGlnLysTyLysSerSerArg-----LysLeuIle 95


```

QY 400 -----ProletrpseralaserProProA 408
DB 4935 GCCGAGTTGGAAGAGTGTCTCCGCGGAGATGCCCTCGACACCGAGAGGCTCTCCA 4876
QY 408 rgaLaProArGserSerThrProCyvProGlyValaValaArgInuAphrThyProV 428
DB 4875 GATCGCTCGACCGAACGTTGAGCCCT----- 4847
QY 428 aGlyThInGInGlyValProSerProAlaLeuAlaGInGlyGlyProGInGlySerTrpA 448
DB 4846 -----CCCGGAGAGCTCTCGGCTGTGCGCGGCGAG----- 4814
QY 448 rGpHeuGInTrpAnSerMetProArGLeuProThraPheuAhpValGInGlyProT 468
DB 4813 -----CCGCTGCACCTTCG----- 4799
QY 468 rPpHeArGHisrYrAhpPheArGInSerCyvTrpValaArgAlaIleSerGInGInuAhpG 488
DB 4799 ----- 4799
QY 488 InLeuAlaProCyvStrpGInaGInuAhpProAlaGInuArgValaArgSerAla-----P 506
DB 4798 --GCAGCGCTTCG-----GATCCAGCTTCGCGTGGCGTTCAGCGGCAAGCT 4753
QY 506 heaLaAlaProSerThraPserAhpArgInGlyThrProPheArGAlaArgAhpGInuAhp 526
DB 4752 GCAGCAGACCTGCGCAAGCGCTGCTGATGACGAGCAGAGCTCGGCGAGCAGCT 4693
QY 526 roCyvAlaProThraPserGlyProCyvLeuCyvGlyLeuHisLeuGInuAhpSerGInuAhp 546
DB 4692 GCGAGCTCCGCGCAGAGAGTCT-----CACCCGAGTACGCTCAGCGCC 4648
QY 546 roPProGly 548
DB 4647 CACGAGT 4640

```

RESULT 15

```

US-09-919-039-243/C
; Sequence 243, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919, 039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-26
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 243
; LENGTH: 3763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Inocyte ID No. 6727066 346209.3
US-09-919-039-243

```

Alignment Scores:

```

Pred. No.: 0.227 Length: 3763
Score: 148.50 Matches: 95
Percent Similarity: 33.87% Conservative: 31
Best Local Similarity: 25.54% Mismatches: 147
Query Match: 4.98% Indels: 99
DB: 4 Gaps: 19

```

US-10-071-838-2 (1-549) x US-09-919-039-243 (1-3763)

```

QY 210 LeuValGInLeuLeuAlaSerGInuArg--HisSerLeuGInGlyPheHis--SerPro 227
DB 1133 CTCCTCAGCTGCTCAGCAGCGCCGCGCTGCTCAGCTCCATGCTCGCAGCAGCCGAGAGCC 1074

```

```

QY 228 AenGlyGlyThrValGInGlyLeuGInuAhpGInGInuAhpValaValaIleArGserGIn 247
DB 1073 CAGGCGCAATTCGCGCGCTGTCTCTCTCGAGAG----- 1038
QY 248 ProLysrHrMetGlyHisGInuAhpLysrValuAhpLeuCyvGlyGInCyvSerProLeuGly 267
DB 1037 -----CTCGGAGATGTCAGCTGGAGCTTC 1011
QY 268 CyvLeuIleArGInLeu-----IleAhpGlyIleSerLeuGlyLeuThrLeuArg--- 284
DB 1010 TGTGTGTCTCAGCTCTTGTATAGAGAGATGTTTGAAGCTCAGCTGTTCAGCGCCCC 951
QY 285 -----LeuThraPheValTyLeuValGInGlyGInuAhpLeu 297
DB 950 AGACACATTCGCAATTCAGAGCTGGGCTCGCTGCAAGCTTCACAGAGCTCGCGTGC 891
QY 298 --MetProIleThraArgIleAlaPheLysrValGInGlyLysrLeu--ThrLysrHse 316
DB 890 CCTTCCCATCAGCAGCGCGGCTCGACCTCCGACAGCTGTATACGAGCGGCGCT 831
QY 316 rArGcyvGlyProTrpAlaArgPheCyvAhpAhpPheValaAhpThrTrpAlaArgAhpG 336
DB 830 CGGCTGCTCAGCCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 774
QY 336 uAhpThrValLeuLysrHisLeu--ArgAlaSerMetLysrLeuThrArgLysrGInGlyA 356
DB 773 GCAGACCTCGGCTCAGAGCGGCGGAGCTTCCTCCATGACCTTCGCGCGGCG 714
QY 356 sPLeuPro-----ProProAlaLysrProGInGly--SerSerAlaSerArGProV 373
DB 713 CTGCTCAGAGCTCTCGCAGAGCTCCCGAGGCGCTGCGGCTGAGCTCTCAGCTCTCG 654
QY 373 aLProAlaSerArGInGlyLysrThrLeuCyvGlyGly-----AhpArgGInaIleAhp 390
DB 653 CAGTCCGCGCTGTGA-----TGGAGTCTCTCTCCAGAGCTCTTCTTCAg 606
QY 390 roPProGlyProProAlaArgPheProArGProIleTrpSerAlaSerPro----- 406
DB 605 CTCCTCAGCTCTCTCT-----CCCTTGAACCGAGACTCTCTGTCGCTG 558
QY 407 -----ProArgAlaProArGserSerThrProCyvP 417
DB 557 GGTGAGTCAAGCGTGTCTCCAGCTCGCGCGAGCGCTCCAGCTCTCGCGAGGTC 498
QY 417 roGlyGlyValaValaArgInuAhpThrTrpPro--ValGlyThrGInGlyValProSerP 436
DB 497 CCGGCGCTGCTTCGCGCTGTGCTGTGCGCACAGCTCAGACTCAGACTCTCTGGGCG 438
QY 436 ro-----AlaLeuAlaGInGlyGlyProGInGlySerT 447
DB 437 CTCGCGCAGGCTGTGAGCTCCCGAGAGATTCAGAGCTGGCGCGCGCCAGCC 378
QY 447 rPArGpHeuGInTrpAnSerMetProArGLeuProThraPheuAhpValGInGlyP 467
DB 377 CTCGCTTCTGCGCTGCGCAGGCGAGCTGAGCTCTCTCTCCGCGCAGCTGGGCG 318
QY 467 roTrpPheArGHisrYrAhpPheArGInSerCyvTrpValaArgAlaIleSerGInGInuA 487
DB 317 CC-----GCAGCTCTCTGCGCGCTGTGCTGCTCCAG 285
QY 487 aPGLnLeuAlaProCyvStrpGInaGInuAhpProAlaGInuArgSerAlaPheA 507
DB 284 CATGCTCTCTGAGCTGTGAGCTCTCCCATCAGC-----TCGCTTCAG 237
QY 507 laaLaProSerThraPserAhpArgInGlyThrProPheArGAlaArgAhpGInuAhpProC 527
DB 236 CTCCTCAGCTCTCTGCG-----GACCTT 213
QY 527 yAaLaProThraSerGly-----ProCyvLeu 535
DB 212 CTCCTCTTCGTAGCGAGTCTCTCATGTCG 181

```

Mon Feb 7 06:58:11 2005

us-10-071-838-2.rml

Page 16

Search completed: February 4, 2005, 11:53:57
Job time : 198 secs

SQ Sequence 549 AA;

Query Match

Best Local Similarity 100.0%; Score 2983; DB 5; Length 549;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDVVEVAGSMAWAOEREDIMKYEKGRAGLPEDKGPFRPSYNNVNDHIGIVHETELPPL 60
DB 1 MDVVEVAGSMAWAOEREDIMKYEKGRAGLPEDKGPFRPSYNNVNDHIGIVHETELPPL 60
QY 61 TAREAKQIRREIRSRKSWMDLGDWEKYSRKLIDRAYKGMPIIRGPMVSLINIEEM 120
DB 61 TAREAKQIRREIRSRKSWMDLGDWEKYSRKLIDRAYKGMPIIRGPMVSLINIEEM 120
QY 121 KLNKPGRYOIMKEKGRSSEHIQRIIDRVSGTLRKIIIFRDRYGTQKRELIHLLAYEY 180
DB 121 KLNKPGRYOIMKEKGRSSEHIQRIIDRVSGTLRKIIIFRDRYGTQKRELIHLLAYEY 180
QY 181 NPEVGYCRDLSHIAALFLYLPEEDAFAVLVOLASERHSIQGHSFPGSGTVQGLDQOE 240
DB 181 NPEVGYCRDLSHIAALFLYLPEEDAFAVLVOLASERHSIQGHSFPGSGTVQGLDQOE 240
QY 241 HVVATSQPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGQALMPI 300
DB 241 HVVATSQPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGQALMPI 300
QY 301 TRIAFVQOKRLTKTSRCGPMARFCNRPVDTWARDDEDTVLKHLRASMKCLTRKQGLPP 360
DB 301 TRIAFVQOKRLTKTSRCGPMARFCNRPVDTWARDDEDTVLKHLRASMKCLTRKQGLPP 360
QY 361 APEEGSSASRPVPASRGKTLCKGDRAPPGPARPPRPRTWSASPPRARSSTPCPGA 420
DB 361 APEEGSSASRPVPASRGKTLCKGDRAPPGPARPPRPRTWSASPPRARSSTPCPGA 420
QY 421 VREDIYVPGTGVSPALAGGPGQSWRFLQNMSPRLPTDLVVEGPMFRHYDFRQSCWV 480
DB 421 VREDIYVPGTGVSPALAGGPGQSWRFLQNMSPRLPTDLVVEGPMFRHYDFRQSCWV 480
QY 481 RAISOEDOLAPCWAQEHFAERVSAPFAASTDSOGTPPRARDEQCAPTSGPCGLGHL 540
DB 481 RAISOEDOLAPCWAQEHFAERVSAPFAASTDSOGTPPRARDEQCAPTSGPCGLGHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549

RESULT 2
ADCC37383
ID ADC37383 standard; protein; 549 AA.
XX
AC ADC37383;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 216.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KW Immunomodulator; Cytostatic; Antimicrobial; Osteopontin; Anti-HIV;
KW Neuroprotective; Noctropic; Cardiant; Gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
```

PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
XX
XX (ASAH) KASEI KK.

XX Matsuda A, Muramatsu S;
XX
XX WPI; 2003-505282/47.
DR N-PSDB; ADC37382.

XX
XX
XX New purified protein that activates nuclear factor kappa B (NF-kappaB),
XX useful for treating inflammation, autoimmune diseases, cancers,
XX infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
XX ischemic disorders.
XX
XX Claim 1, SEQ ID NO 216; 938bp; English.

XX
XX The present invention relates to novel proteins and their coding
XX sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
XX kappaB). The proteins and their coding sequences are useful for treating
XX a disease associated with NF-kappaB activation, such as inflammation,
XX autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
XX neurodegenerative diseases, or ischemic disorders.

SQ Sequence 549 AA;

Query Match 100.0%; Score 2983; DB 7; Length 549;
Best Local Similarity 100.0%; Pred. No. 1.4e-257;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDVVEVAGSMAWAOEREDIMKYEKGRAGLPEDKGPFRPSYNNVNDHIGIVHETELPPL 60
DB 1 MDVVEVAGSMAWAOEREDIMKYEKGRAGLPEDKGPFRPSYNNVNDHIGIVHETELPPL 60
QY 61 TAREAKQIRREIRSRKSWMDLGDWEKYSRKLIDRAYKGMPIIRGPMVSLINIEEM 120
DB 61 TAREAKQIRREIRSRKSWMDLGDWEKYSRKLIDRAYKGMPIIRGPMVSLINIEEM 120
QY 121 KLNKPGRYOIMKEKGRSSEHIQRIIDRVSGTLRKIIIFRDRYGTQKRELIHLLAYEY 180
DB 121 KLNKPGRYOIMKEKGRSSEHIQRIIDRVSGTLRKIIIFRDRYGTQKRELIHLLAYEY 180
QY 181 NPEVGYCRDLSHIAALFLYLPEEDAFAVLVOLASERHSIQGHSFPGSGTVQGLDQOE 240
DB 181 NPEVGYCRDLSHIAALFLYLPEEDAFAVLVOLASERHSIQGHSFPGSGTVQGLDQOE 240
QY 241 HVVATSQPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGQALMPI 300
DB 241 HVVATSQPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGQALMPI 300
QY 301 TRIAFVQOKRLTKTSRCGPMARFCNRPVDTWARDDEDTVLKHLRASMKCLTRKQGLPP 360
DB 301 TRIAFVQOKRLTKTSRCGPMARFCNRPVDTWARDDEDTVLKHLRASMKCLTRKQGLPP 360
QY 361 APEEGSSASRPVPASRGKTLCKGDRAPPGPARPPRPRTWSASPPRARSSTPCPGA 420
DB 361 APEEGSSASRPVPASRGKTLCKGDRAPPGPARPPRPRTWSASPPRARSSTPCPGA 420
QY 421 VREDIYVPGTGVSPALAGGPGQSWRFLQNMSPRLPTDLVVEGPMFRHYDFRQSCWV 480
DB 421 VREDIYVPGTGVSPALAGGPGQSWRFLQNMSPRLPTDLVVEGPMFRHYDFRQSCWV 480
QY 481 RAISOEDOLAPCWAQEHFAERVSAPFAASTDSOGTPPRARDEQCAPTSGPCGLGHL 540
DB 481 RAISOEDOLAPCWAQEHFAERVSAPFAASTDSOGTPPRARDEQCAPTSGPCGLGHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549
```

RESULT 3
NAT84901

XX	AAV84901 standard; protein; 549 AA.
AC	AAV84901;
D7	21-AUG-2000 (first entry)
DE	A human proliferation and apoptosis related protein.
KW	Human; proliferation and apoptosis related protein; PROAP; porriasis; cell proliferative disorder; immunological disorder; hepatitis; reproductive disorder; arteriosclerosis; cirrhosis; melanoma; lymphoma; cancer; acquired immune deficiency syndrome; AIDS; allergy; anaemia; asthma; diabetes mellitus; osteoarthritis; endometriosis; uterine fibroid; menstrual cycle.
OS	Homo sapiens.
FT	Key
FT	Modified-site
FT	/note= "potential phosphorylation site"
FT	73
FT	/note= "potential phosphorylation site"
FT	90
FT	/note= "potential phosphorylation site"
FT	91
FT	/note= "potential phosphorylation site"
FT	98..315
FT	/note= "probable rabgap domain"
FT	99
FT	/note= "potential phosphorylation site"
FT	139
FT	/note= "potential phosphorylation site"
FT	152
FT	/note= "potential phosphorylation site"
FT	216
FT	/note= "potential phosphorylation site"
FT	282
FT	/note= "potential phosphorylation site"
FT	313
FT	/note= "potential phosphorylation site"
FT	315
FT	/note= "potential phosphorylation site"
FT	346
FT	/note= "potential phosphorylation site"
FT	351
FT	/note= "potential phosphorylation site"
FT	446
FT	/note= "potential phosphorylation site"
FT	460
FT	/note= "potential phosphorylation site"
FT	484
FT	/note= "potential phosphorylation site"
FT	511
FT	/note= "potential phosphorylation site"
XX	WO200023589-A2.
PN	27-APR-2000.
PD	19-OCT-1999;
PF	99WO-US024511.
XX	20-OCT-1998;
PR	98US-017221P.
PR	04-FEB-1999;
PR	99US-011855P.
PR	11-FEB-1999;
PR	99US-017222P.
PR	22-APR-1999;
PR	99US-015433P.
PA	(INCY-) INCYTE PHARM INC.
XX	Tang YT, Yue H, Hillman JL, Guegler KJ, Corley NC, Lal P;
P1	Azimzal Y, Baughn MR, Yang J, Shin L;
XX	WP1; 2000-339686/29.
DR	N-PSDB; AAA15001.

Remained
Empty

XX	New human proliferation and apoptosis related protein polypeptides used
PT	for diagnosis, treatment and prevention of cell proliferative,
PT	immunological and reproductive disorders.
XX	
PS	Claim 1; Page 101-102; 128BP; English.
XX	
CC	The present sequence represents a human proliferation and apoptosis
CC	related protein (PROAP). The polypeptides and polynucleotides can be used
CC	for the diagnosis, treatment and prevention of cell proliferative,
CC	immunological and reproductive disorders. Disorders associated with
CC	decreased expression or activity of include arteriosclerosis, cirrhosis,
CC	hepatitis, psoriasis, melanoma, lymphoma and cancers of the breast, brain
CC	and prostate, acquired immune deficiency syndrome (AIDS), allergies,
CC	anaemia, asthma, diabetes mellitus, osteoarthritis, endometriosis,
CC	uterine fibroids and disruptions of the menstrual cycle. Antibodies
CC	against PROAP can be used in diagnosis of disorders characterized by PROAP
CC	e.g., in ELISA (enzyme linked immunosorbent assays) and the
CC	polymicrotiter arrays may be used to detect and quantify gene expression in
CC	biopsied tissues. These techniques can also be used to monitor regulation
CC	of PROAP levels during therapeutic intervention
XX	
SQ	Sequence 549 AA;
	Query Match 99.8%; Score 2978; DB 3; Length 549;
	Best Local Similarity 99.8%; Pred. No. 3.9e-257;
	Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 MDVVEVAGSWMWOEREDIMKYEKGHAGLPEDKPKPPRSYNNNVNDLGIHVHTEPLPPL 60
DG	1 MDVVEVAGSWMWOEREDIMKYEKGHAGLPEDKPKPPRSYNNNVNDLGIHVHTEPLPPL 60
OY	61 TAREAKOIRREISRKSKFWVDMLGDWEYKSSRKLIDRAYKGMPIINRGPMWSVLNITEEM 120
DG	61 TAREAKOIRREISRKSKFWVDMLGDWEYKSSRKLIDRAYKGMPIINRGPMWSVLNITEEM 120
OY	121 KLRKPGRYOVMKEGKRSSSHIQRIDVDSTGLRKHIFFPDROYTKORELLHILAAEEY 180
DG	121 KLRKPGRYOVMKEGKRSSSHIQRIDVDSTGLRKHIFFPDROYTKORELLHILAAEEY 180
OY	181 NPEVGCRYDLSHTNALFLYLYPEDEAFMALVOLLASRRHSLOGFHSNPGGTVOGLDOQE 240
DG	181 NPEVGCRYDLSHTNALFLYLYPEDEAFMALVOLLASRRHSLOGFHSNPGGTVOGLDOQE 240
OY	241 HVVATSOPTMGHDKKDLGCQSPLCLRLILDIGISLGITRLMDVLYVEBQALMPI 300
DG	241 HVVATSOPTMGHDKKDLGCQSPLCLRLILDIGISLGITRLMDVLYVEBQALMPI 300
OY	301 TRIAFKOQXRLTKTSRCGPWFARCNFVDTWADBEDTVLKHLRASMKKLTTRKGGDLP 360
DG	301 TRIAFKOQXRLTKTSRCGPWFARCNFVDTWADBEDTVLKHLRASMKKLTTRKGGDLP 360
OY	361 AKPQGSSASRPVVASGCTLKCKGDQAAPGARPRPRTWSAPRARASSSPPCGA 420
DG	361 AKPQGSSASRPVVASGCTLKCKGDQAAPGARPRPRTWSAPRARASSSPPCGA 420
OY	421 VREBTVTVGTQGVNSPALAQSGPQSWRFLOMNSMPRLPTDLLDVGEFMFHAYDFROSQV 480
DG	421 VREBTVTVGTQGVNSPALAQSGPQSWRFLOMNSMPRLPTDLLDVGEFMFHAYDFROSQV 480
OY	481 RAISOEBOLAPCWAHEPAERVNSAFAPASTDSOGTTPRARDEQPACAPISGCLGLHL 540
DG	481 RAISOEBOLAPCWAHEPAERVNSAFAPASTDSOGTTPRARDEQPACAPISGCLGLHL 540
OY	541 ESSGFPPGF 549
DG	541 ESSGFPPGF 549

AC ADC37385;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 218.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischemic disorder; Anti-inflammatory;
KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
XX
OS Homo sapiens.
XX
PN MO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002MO-JP012644.
XX
PR 03-DEC-2001; 2001JP-00368692.
PR 05-DEC-2001; 2001US-0335829P.
PR 03-OCT-2002; 2002JP-00291302.
PR 04-OCT-2002; 2002US-0415769P.
XX
PA (ASAH) ASAH KASEI KK.
XX
PI Matsuda A, Muramatsu S;
XX
DR WPI; 2003-505282/47.
XX
DR N-PSDB; ADC37384.
XX
PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
PT useful for treating inflammation, autoimmune diseases, cancers,
PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
PT ischemic disorders.
XX
PS Claim 1; SEQ ID NO 218; 938bp; English.
XX
CC The present invention relates to novel proteins and their coding
CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
CC kappaB). The proteins and their coding sequences are useful for treating
CC a disease associated with NF-kappaB activation, such as inflammation,
CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
CC neurodegenerative diseases, or ischemic disorders.
XX
SQ Sequence 549 AA;
Query Match 99.7%; Score 2975; DB 7; Length 549;
Best Local Similarity 99.8%; Pred. No. 7.2e-257;
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MDVVEVAGSWMAQEREDIIIMKYKGRAGLPEDKGPFRSYNNVNDHGIYHETELPPL 60
DB 1 MDVVEVAGSWMAQEREDIIIMKYKGRAGLPEDKGPFRSYNNVNDHGIYHETELPPL 60
QY 61 TAREBAQIIRREISRKSKWMDLGMWYKSSRKLIDRAYKGMPIIRGPMVSLNITEM 120
DB 61 TAREBAQIIRREISRKSKWMDLGMWYKSSRKLIDRAYKGMPIIRGPMVSLNITEM 120
QY 121 KLNKPGRYOIMKEKGRSEHIORIDRVSGTLRKHIFFRDRYGTQRELLHILLAYEY 180
DB 121 KLNKPGRYOIMKEKGRSEHIORIDRVSGTLRKHIFFRDRYGTQRELLHILLAYEY 180
QY 181 NBEVGYCRDLISHIALFLTYLPEDAFWALVOLASERHSIQGFHSPMGTVQGLQDOOE 240
DB 181 NBEVGYCRDLISHIALFLTYLPEDAFWALVOLASERHSIQGFHSPMGTVQGLQDOOE 240
QY 241 HVVATSQPTMGHOKKDI CGCQSPGLCTIRLLIDGSLGTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSQPTMGHOKKDI CGCQSPGLCTIRLLIDGSLGTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVOQKRLTTSRCGPMARFCNRFPVDMARDSDIVLGLHLSAKMKLTKKQGLDPPP 360

DB 301 TRIAFKVOQKRLTTSRCGPMARFCNRFPVDMARDSDIVLGLHLSAKMKLTKKQGLDPP 360
QY 361 AKPEQSSASRPVPAIRGGTLCCKGRQAPPPGPAPFPPIPIVSASPAPRSSTPCPGA 420
DB 361 AKPEQSSASRPVPAIRGGTLCCKGRQAPPPGPAPFPPIPIVSASPAPRSSTPCPGA 420
QY 421 VREDTYPVGTQVSPALAQGGPQGSWRFLQNSMRLPTDLVDVGPMPRHYPFGOSCMV 480
DB 421 VREDTYPVGTQVSPALAQGGPQGSWRFLQNSMRLPTDLVDVGPMPRHYPFGOSCMV 480
QY 481 RAISOEDOLAPCQMAHPEARVRSAPAPSTDSQCTPPRARDQPCAPTSGLCGLHT 540
DB 481 RAISOEDOLAPCQMAHPEARVRSAPAPSTDSQCTPPRARDQPCAPTSGLCGLHT 540
QY 541 ESSQPPPGP 549
DB 541 ESSQPPPGP 549
RESULT 5
ABUS3234
ID ABUS3234 standard; protein; 549 AA.
XX
AC ABUS3234;
XX
DT 14-APR-2003 (first entry)
XX
DE Human cell cycle-associated protein from DKFzphb23_35p22.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN MO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000MO-IB001496.
XX
PR 18-AUG-1999; 99US-0149499P.
PR 28-SEP-1999; 99US-0156503P.
XX
PA (GERH-) GERMAN HUMAN GENOME PROJECT.
XX
PI Wiemann S;
XX
DR WPI; 2001-327840/34.
DR N-PSDB; ABX71403.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of
PT different human tissues, useful in recombinant DNA methodologies.
XX
PS Claim 21; Page 868; 1095pp; English.
XX
CC This invention describes novel polynucleotides and polypeptides isolated
CC from human cDNA libraries which can be used for gene therapy or in
CC vaccines. The polynucleotides of the invention and antibodies encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The products of the
CC invention may also be used to identify modulators of expression and
CC the activity and to down regulate expression and activity. The antibodies of
CC the invention may also be used as diagnostic agents for detecting the
CC presence of polypeptides in samples. This sequence represents a
CC polypeptide described in the disclosure of the invention
XX
SQ Sequence 549 AA;
Query Match 99.3%; Score 2963; DB 4; Length 549;
Best Local Similarity 99.3%; Pred. No. 8.5e-256;
Matches 545; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDVVEVAGSWMAQEREDIIIMKYKGRAGLPEDKGPFRSYNNVNDHGIYHETELPPL 60

```

Db      1 MOVVEVAGSWMAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGIHETELPPL 60
Qy      61 TAREAKQIRREISRSKSKVDMGDMWEKYSRKILDRAYKMPNIRGPMWSVLNITEEM 120
Db      61 TAREAKQIRREISRSKSKVDMGDMWEKYSRKILDRAYKMPNIRGPMWSVLNITEEM 120
Qy      121 KLNKNGRQIOMEKKGKSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQORELHILAYEEY 180
Db      121 KLNKNGRQIOMEKKGKSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQORELHILAYEEY 180
Qy      181 NPEVGYCRDLSHIALFLYLPEBEDAFMALVOLLASERHSLOGFHSFNGGTVOGLDQOE 240
Db      181 NPEVGYCRDLSHIALFLYLPEBEDAFMALVOLLASERHSLOGFHSFNGGTVOGLDQOE 240
Qy      241 HVVATSOPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOLAMP 300
Db      241 HVVATSOPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOLAMP 300
Qy      301 TRIAFKVOQKRLTKTSRGGPMARFCNRFVDTWARDEDTVLKHLRASMKGKLRKOGDLPPP 360
Db      301 TRIAFKVOQKRLTKTSRGGPMARFCNRFVDTWARDEDTVLKHLRASMKGKLRKOGDLPPP 360
Qy      361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSTPCPGGA 420
Db      361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSTPCPGGA 420
Qy      421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOMNSMPLPTDLVDEGPMFRHYDFROSCW 480
Db      421 VREDTYPVGTQGVSPALAQGGPQGSWRFLOMNSMPLPTDLVDEGPMFRHYDFROSCW 480
Qy      481 PAISQEDOLAPCWOAHEPAERVSAPAAPSTSDQTPRARDEQCAPTSGPCGGLH 540
Db      481 PAISQEDOLAPCWOAHEPAERVSAPAAPSTSDQTPRARDEQCAPTSGPCGGLH 540
Qy      541 ESSQPPPGF 549
Db      541 ESSQPPPGF 549

RESULT 6
ABG70737
ID ABG70737 standard; protein; 610 AA.
AC ABG70737;
XX
XX 28-NOV-2002 (first entry)
XX
XX Human PRC17 protein splice variant 1.
DE
XX Human, PRC17; prostate cancer; ovarian cancer; lung cancer;
KW breast cancer; cytosolic; chromosome 17q11-12; splice variant.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX MISC-difference 141
XX FT /note= "Encoded by AT"
XX
XX WO200262958-A2.
XX
XX 15-AUG-2002.
XX
XX 08-FEB-2002; 2002WO-US003457.
XX
XX 08-FEB-2001; 2001US-0267615P.
XX
XX (TUL- ) TULARIK INC.
XX
XX LI J, Powers S, Xiang P, Peng Y;
XX
XX WPI; 2002-706902/76.
XX
XX N-PsDB; AB854707.

```

```

XX
XX Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PT or expression relative to normal, for example cancer.
XX
XX Claim 2; Page 64; 78pp; English.
XX
XX The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present amino acid sequence represents the human PRC17
CC protein splice variant 1 of the invention. The human PRC17 gene is
CC located on chromosome 17q11-12
XX
XX Sequence 610 AA;
SQ
Query Match 98.4%; Score 2934.5; DB 5; Length 610;
Best Local Similarity 89.8%; Pred. No. 3.5e-253;
Matches 548; Conservative 0; Mismatches 1; Indels 61; Gaps 1;
Qy      1 MOVVEVAGSWMAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGIH----- 53
Db      1 MOVVEVAGSWMAQEREDIMKYEKGHRAGLPEDKGPFRSYNNVNDHLGIHSCRSWMS 60
Qy      54 -----ETELPP 59
Db      61 APQEGPCPPFPVSPGLSPELERDRASPFWGSAAPRLGPLQAPCSSALPGLPYSETELP 120
Qy      60 LTRAKQIRREISRSKSKVDMGDMWEKYSRKILDRAYKMPNIRGPMWSVLNITEE 119
Db      121 LTRAKQIRREISRSKSKVDMGDMWEKYSRKILDRAYKMPNIRGPMWSVLNITEE 180
Qy      120 MCLKNPGRQIOMEKKGKSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQORELHILAYEE 179
Db      181 MCLKNPGRQIOMEKKGKSSSEHIQRIIDRVSGTLRKHIIFRDRYGTQORELHILAYEE 240
Qy      180 YNPEVGYCRDLSHIALFLYLPEBEDAFMALVOLLASERHSLOGFHSFNGGTVOGLDQO 239
Db      241 YNPEVGYCRDLSHIALFLYLPEBEDAFMALVOLLASERHSLOGFHSFNGGTVOGLDQO 300
Qy      240 EHVATSOPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOLAMP 299
Db      301 EHVATSOPTMGHODKDLGCGCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOLAMP 360
Qy      300 ITRIAFKVOQKRLTKTSRGGPMARFCNRFVDTWARDEDTVLKHLRASMKGKLRKOGDLPP 359
Db      361 ITRIAFKVOQKRLTKTSRGGPMARFCNRFVDTWARDEDTVLKHLRASMKGKLRKOGDLPP 420
Qy      360 PAKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSTPCPGGA 419
Db      421 PAKPEQSSASRPVPASRGKTLCKGDRQAPPGPARFPPIWSASPPRAPRSTPCPGGA 480
Qy      420 AVREDTYPVGTQGVSPALAQGGPQGSWRFLOMNSMPLPTDLVDEGPMFRHYDFROSCW 479
Db      481 AVREDTYPVGTQGVSPALAQGGPQGSWRFLOMNSMPLPTDLVDEGPMFRHYDFROSCW 540
Qy      480 VRAISOEDOLAPCWOAHEPAERVSAPAAPSTSDQTPRARDEQCAPTSGPCGGLH 539
Db      541 VRAISOEDOLAPCWOAHEPAERVSAPAAPSTSDQTPRARDEQCAPTSGPCGGLH 600
Qy      540 LESSQPPPGF 549
Db      601 LESSQPPPGF 610

RESULT 7
ABP69593
ID ABP69593 standard; protein; 531 AA.

```

XX AC ABP69593;
 XX DT 20-JAN-2003 (first entry)
 XX DE Human polypeptide SEQ ID NO 1640.
 XX
 XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulnerrary; fungicide; antibacterial; virucide; protozoacide;
 KW antiarthritic.
 XX
 XX Homo sapiens.
 XX OS
 XX MO200270539-A2.
 XX
 XX PD 12-SEP-2002.
 XX
 XX PF 05-MAR-2002; 2002WO-US005095.
 XX
 XX PR 05-MAR-2001; 2001US-00799451.
 XX
 XX PA (HYSE-) HYSEQ INC.
 XX
 XX PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F,
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 XX DR WPI; 2002-759812/82.
 XX
 XX DR N-PSDB; AB211810.
 XX
 XX PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 PT
 XX
 XX PS Claim 9; SEQ ID NO 1640; 1012pp + Sequence listing; English.
 XX
 XX CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (AB211119-
 CC AB212066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, wound, burns, incision, ulcers, liver
 CC platelet or coagulation disorders, viral, fungal, parasitic), liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at fip.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 531 AA;
 XX
 XX Query Match 95.5%; Score 2850; DB 5; Length 531;
 XX Best Local Similarity 96.2%; Pred. No. 1e-245;
 XX Matches 528; Conservative 1; Mismatches 2; Indels 18; Gaps 1;
 XX
 XX QY 1 MDVVAAGSWMAQOEEDIIIMKYEKGRAGLPEDKPKPRFSYNNVNLGIYHETETPL 60
 XX DB 1 MDVVAAGSWMAQOEEDIIIMKYEKGRAGLPEDKPKPRFSYNNVNLGIYHETETPL 60
 XX QY 61 TAREAKQIRREISRSKSKVMDLGMWEKKSKSKRIDRAYGKMPNIRGPMWSVILNIEM 120
 XX DB 61 TAREAKQIRREISRSKSKVMDLGMWEKKSKSKRIDRAYGKMPNIRGPMWSVILNIEM 120

QY 121 KLANPGRYQIMKEKGRKSSHHIORIDRVSGTLRKHIFFRDRTGKORRELLILAYERY 180
 DB 121 KLANPGRYQIMKEKGRKSSHHIORIDRVSGTLRKHIFFRDRTGKORRELLILAYERY 180
 QY 181 NPEVGYCRDLSHIAALFLYLTPREDAFMALVOLLASERHSLOGFHSPNGVTQSLQDOOE 240
 DB 163 NPEVGYCRDLSHIAALFLYLTPREDAFMALVOLLASERHSLOGFHSPNGVTQSLQDOOE 222
 QY 241 HVAATSQPKTMGHQDKKDLGQCSPLGCLIRILIDIGISLGLRLMDVYLVEGQALMPI 300
 DB 223 HVAATSQPKTMGHQDKKDLGQCSPLGCLIRILIDIGISLGLRLMDVYLVEGQALMPI 282
 QY 301 TRIAFVQOKRLTKTSSRCQPMARFCNRPYDTWABEDTVLKHLPASMKLTKKQDGLPP 360
 DB 283 TRIAFVQOKRLTKTSSRCQPMARFCNRPYDTWABEDTVLKHLPASMKLTKKQDGLPP 342
 QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPGPPAPRPPIWASPPRAPSSTPCPGGA 420
 DB 343 AKPEQSSASRPVPASRGKTLCKGDRQAPGPPAPRPPIWASPPRAPSSTPCPGGA 402
 QY 421 VREDTYPVGTQVSPALAOGSPQSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCWV 480
 DB 403 VREDTYPVGTQVSPALAOGSPQSWRFLOWNSMPRLPTDLVEGPMFRHYDFRQSCWV 462
 QY 481 RAISOEDQLAPCWOAHEPABERVRSAFAAPSTDSOGTPFARAPBOPCAPTSRGCLGLHL 540
 DB 463 RAISOEDQLAPCWOAHEPABERVRSAFAAPSTDSOGTPFARAPBOPCAPTSRGCLGLHL 522
 QY 541 ESSQPPPGF 549
 DB 523 ESSQPPPGF 531
 XX
 XX RESULT 8
 XX ADC37389
 XX ID ADC37389 standard; protein; 527 AA.
 XX
 XX AC ADC37389;
 XX
 XX DT 18-DEC-2003 (first entry)
 XX
 XX DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 222.
 XX
 XX KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.
 XX
 XX OS Homo sapiens.
 XX
 XX PN MO2003048202-A2.
 XX
 XX PD 12-JUN-2003.
 XX
 XX PF 03-DEC-2002; 2002WO-JP012644.
 XX
 XX PR 03-DEC-2001; 2001JP-00368692.
 XX PR 05-DEC-2001; 2001US-0335829P.
 XX PR 03-OCT-2002; 2002JP-00291302.
 XX PR 04-OCT-2002; 2002US-0415769P.
 XX
 XX PA (ASAH) ASAH KASEI KK.
 XX
 XX PI Matsuda A, Muramatsu S;
 XX
 XX DR WPI; 2003-505282/47.
 XX DR N-PSDB; ADC37388.
 XX
 XX PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.

XX PS Claim 1; SEQ ID NO 222; 938bp; English.

CC CC The present invention relates to novel proteins and their coding sequences (ADG37168-ADG37455), which activate nuclear factor kappa B (NF-kappaB). The proteins and their coding sequences are useful for treating a disease associated with NF-kappaB activation, such as inflammation, CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, CC neurodegenerative diseases, or ischemic disorders.

XX SQ Sequence 527 AA;

Query Match 95.0%; Score 2835; DB 7; Length 527;
 Best Local Similarity 95.8%; Pred. No. 2,3e-243; Indels 22; Gaps 1;
 Matches 526; Conservative 0; Mismatches 1;

QY 1 MDVEVAGSWMAQEREDIMKTEKGRAGLPEDKGPFRSYNNVNDHIGIVHETELPPL 60
 DB 1 MDVEVAGSWMAQEREDIMKTEKGRAGLPEDKGPFRSYNNVNDHIGIVHETELPPL 60

QY 61 TAREAKQIRREISRSKSWVDMLGDWEKYSRKLDRAVKGPMNIRGPMWSVLNTEEM 120
 DB 61 TAREAKQIRREISRSKSWVDMLGDWEKYSRKLDRAVKGPMNIRGPMWSVLNTEEM 120

QY 121 KLNKPGRYOIMKEKGRSEHIQRIIDRVSGTLRKHIIFPRDRYGTGKRELIHILAYEY 180
 DB 121 KLNKPGRYOIMKEKGRSEHIQRIIDRVSGTLRKHIIFPRDRYGTGKRELIHILAYEY 180

QY 121 KLNKPGRYOIMKEKGRSEHIQRIIDRVSGTLRKHIIFPRDRYGTGKRELIHILAYEY 180
 DB 121 KLNKPGRYOIMKEKGRSEHIQRIIDRVSGTLRKHIIFPRDRYGTGKRELIHILAYEY 180

QY 181 NBEVGYCRDLSHIALFLLYLPEEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240
 DB 181 NBEVGYCRDLSHIALFLLYLPEEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240

QY 241 HVAATSOPTKMGHODKDLGGCSPGLCIRILINDISGLTLRLMDVYLVEGEQALMPI 300
 DB 241 HVAATSOPTKMGHODKDLGGCSPGLCIRILINDISGLTLRLMDVYLVEGEQALMPI 300

QY 241 HVAATSOPTKMGHODKDLGGCSPGLCIRILINDISGLTLRLMDVYLVEGEQALMPI 278
 DB 241 HVAATSOPTKMGHODKDLGGCSPGLCIRILINDISGLTLRLMDVYLVEGEQALMPI 278

QY 301 TRIAKVQOKRLTKTSRCGPMARFCNRFVDTWARDEDVTLKHLRASMKKLTKQGDLPPP 360
 DB 301 TRIAKVQOKRLTKTSRCGPMARFCNRFVDTWARDEDVTLKHLRASMKKLTKQGDLPPP 360

QY 361 AKPEGSSASRPVPSASRGKTKLCKGRQAPPGPPAPRPPIWSASPPAPRSTPCPGA 420
 DB 361 AKPEGSSASRPVPSASRGKTKLCKGRQAPPGPPAPRPPIWSASPPAPRSTPCPGA 420

QY 339 AKPEGSSASRPVPSASRGKTKLCKGRQAPPGPPAPRPPIWSASPPAPRSTPCPGA 398
 DB 339 AKPEGSSASRPVPSASRGKTKLCKGRQAPPGPPAPRPPIWSASPPAPRSTPCPGA 398

QY 421 VAEETYPVGTQGVSPALAQGGPQGSWRLQNNSMRLELTDLVGEPMFRHYDFRQSCV 480
 DB 421 VAEETYPVGTQGVSPALAQGGPQGSWRLQNNSMRLELTDLVGEPMFRHYDFRQSCV 480

QY 399 VAEETYPVGTQGVSPALAQGGPQGSWRLQNNSMRLELTDLVGEPMFRHYDFRQSCV 458
 DB 399 VAEETYPVGTQGVSPALAQGGPQGSWRLQNNSMRLELTDLVGEPMFRHYDFRQSCV 458

QY 481 RAISOEDOLAPCMQAEHPAERVRSAPAASTSDSGTPRARDEOCAPTSGPCLCGHL 540
 DB 481 RAISOEDOLAPCMQAEHPAERVRSAPAASTSDSGTPRARDEOCAPTSGPCLCGHL 540

QY 459 RAISOEDOLAPCMQAEHPAERVRSAPAASTSDSGTPRARDEOCAPTSGPCLCGHL 518
 DB 459 RAISOEDOLAPCMQAEHPAERVRSAPAASTSDSGTPRARDEOCAPTSGPCLCGHL 518

QY 541 ESSQEPFPGF 549
 DB 541 ESSQEPFPGF 527

RESULT 9
 ABP69592
 ID ABP69592 standard; protein; 527 AA.
 AC ABP69592;
 XX
 XX 20-JAN-2003 (first entry)
 DE Human polypeptide SEQ ID NO 1639.
 XX
 XX Human, genome mapping; gene therapy; food supplement; virus; fungus;
 KM cell-proliferative disorder; neurodegenerative disease; bacterial;
 KM Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KM multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KM arthritis; cystostatic; immunomodulator; nootropic; neuroprotective;
 KM antiparkinsonian; antidiabetic; immunosuppressive; dermatological;

KM haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
 KM antidiabetic.

XX OS Homo sapiens.
 XX PN WO200270539-A2.
 XX 12-SEP-2002.
 XX 05-MAR-2002; 2002WO-US005095.
 XX 05-MAR-2001; 2001US-00799451.
 XX (HGSB-) HSEQ INC.
 XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QH, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehman T, Wang J, Wang D, Drmanac RT;
 DR MPI; 2002-759812/82.
 DR N-PSDB; AB211809.
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.

PS Claim 9; SEQ ID NO 1639; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (1) comprising a
 CC nucleotide sequence selected from any of 948 sequences (AB211119-
 CC AB212066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ABP6902-ABP6949) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at fcp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 527 AA;

Query Match 94.8%; Score 2827; DB 5; Length 527;
 Best Local Similarity 95.6%; Pred. No. 1.2e-243; Indels 22; Gaps 1;
 Matches 525; Conservative 0; Mismatches 2;

QY 1 MDVEVAGSWMAQEREDIMKTEKGRAGLPEDKGPFRSYNNVNDHIGIVHETELPPL 60
 DB 1 MDVEVAGSWMAQEREDIMKTEKGRAGLPEDKGPFRSYNNVNDHIGIVHETELPPL 60

QY 61 TAREAKQIRREISRSKSWVDMLGDWEKYSRKLDRAVKGPMNIRGPMWSVLNTEEM 120
 DB 61 TAREAKQIRREISRSKSWVDMLGDWEKYSRKLDRAVKGPMNIRGPMWSVLNTEEM 120

QY 121 KLNKPGRYOIMKEKGRSEHIQRIIDRVSGTLRKHIIFPRDRYGTGKRELIHILAYEY 180
 DB 121 KLNKPGRYOIMKEKGRSEHIQRIIDRVSGTLRKHIIFPRDRYGTGKRELIHILAYEY 180

QY 181 NBEVGYCRDLSHIALFLLYLPEEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240
 DB 181 NBEVGYCRDLSHIALFLLYLPEEDAFMALVOLLASERHSLOGFHSPNGTVQGLDQOE 240

QY 241 HVAATSOPTKMGHODKDLGGCSPGLCIRILINDISGLTLRLMDVYLVEGEQALMPI 300
 DB 241 HVAATSOPTKMGHODKDLGGCSPGLCIRILINDISGLTLRLMDVYLVEGEQALMPI 300

QY 301 TRIAKVQOKRLTKTSRCGPMARFCNRFVDTWARDEDVTLKHLRASMKKLTKQGDLPPP 360
 DB 301 TRIAKVQOKRLTKTSRCGPMARFCNRFVDTWARDEDVTLKHLRASMKKLTKQGDLPPP 360

```

Db      279  TRIAFKVOQKRLTITSRGPMARFCNRFVDTWARDEDTYVKKHLRASKMKLTRKOGDIQPP 338
Qy      361  AKPEGGSSASRPVPASRGKTLCKGDRQAPPGPARPPRPIMWSAPPRAPRSSTPCGGA 420
Db      339  AKPEGGSSASRPVPASRGKTLCKGDRQAPPGPARPPRPIMWSAPPRAPRSSTPCGGA 398
Qy      421  VREDTYPVGTGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVGEPMFRHYDFRQSCWV 480
Db      399  VREDTYPVGTGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVGEPMFRHYDFRQSCWV 458
Qy      481  RAISOEDQLAPCQWQAEHPARVRSAPFAAPSTDSDQGTFFRARDPOCAPTSGPCLCGHL 540
Db      459  RAISOEDQLAPCQWQAEHPARVRSAPFAAPSTDSDQGTFFRARDPOCAPTSGPCLCGHL 518
Qy      541  ESSQPPPGF 549
Db      519  ESSQPPPGF 527

```

RESULT 10

ADCC37387 ID ADC37387 standard; protein; 527 AA.

AC ADC37387;

DT 18-DEC-2003 (first entry)

DE Nuclear factor kappa B (NF-kappaB) activating protein, SEQ ID 220.

KM Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;

KW cancer; infectious disease; bone disease; AIDS;

KW neurodegenerative disease; ischaemic disorder; Antinflammatory;

KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;

KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human.

OS Homo sapiens.

PN W02003048202-A2.

PD 12-JUN-2003.

PF 03-DEC-2002; 2002WO-JP012644.

PR 03-DEC-2001; 2001JP-00368692.

PR 05-DEC-2001; 2001US-0335829P.

PR 03-OCT-2002; 2002JP-00291302.

PR 04-OCT-2002; 2002US-0415769P.

PA (ASAH) ASAH KASEI KK.

PI Matsuda A, Muramatsu S;

DR WPI; 2003-505282/47.

DR N-PSDB; ADC37386.

PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 useful for treating inflammation, autoimmune diseases, cancers,
 infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 ischemic disorders.

PS Claim 1; SEQ ID NO 220; 938bp; English.

CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.

SO Sequence 527 AA;

Query Match 94.8%; Score 2827; DB 7; Length 527;

Best Local Similarity 95.6%; Pred. No. 1.2e-243;
 Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;

```

Qy      1  MDVVEVAGSWMAQERDDIIMKTEKGRAGLPEDKPKPPRSYNNVDHLGIYHETELPPL 60
Db      1  MDVVEVAGSWMAQERDDIIMKTEKGRAGLPEDKPKPPRSYNNVDHLGIYHETELPPL 60
Qy      61  TAREAQIIRREISRSKSNWVMDLGDWEKYSRKLDRAYKGMPMNIRGPMWSVLLNTEEM 120
Db      61  TAREAQIIRREISRSKSNWVMDLGDWEKYSRKLDRAYKGMPMNIRGPMWSVLLNTEEM 120
Qy      121  KAKNRRYVIMKEKGRSESHIQRIDRVSGTLRRGHIFRDRYGTGQRELHILLAYEY 180
Db      121  KAKNRRYVIMKEKGRSESHIQRIDRVSGTLRRGHIFRDRYGTGQRELHILLAYEY 180
Qy      181  NPEVGYCRDLSHIALFLTYLPEDAFVALVOLLAERHSIQFHSPPNGTVQGLQDOQE 240
Db      181  NPEVGYCRDLSHIALFLTYLPEDAFVALVOLLAERHSIQFHSPPNGTVQGLQDOQE 240
Qy      241  HVVATSQPKTWGHDQDKDLQCGQSPGLCLIRLLIDGISGLTLRLAMPVYLVEGQALMPI 300
Db      241  HVVATSQPKTWGHDQDKDLQCGQSPGLCLIRLLIDGISGLTLRLAMPVYLVEGQALMPI 278
Qy      301  TRIAFKVOQKRLTITSRGPMARFCNRFVDTWARDEDTYVKKHLRASKMKLTRKOGDIQPP 360
Db      279  TRIAFKVOQKRLTITSRGPMARFCNRFVDTWARDEDTYVKKHLRASKMKLTRKOGDIQPP 338
Qy      361  AKPEGGSSASRPVPASRGKTLCKGDRQAPPGPARPPRPIMWSAPPRAPRSSTPCGGA 420
Db      339  AKPEGGSSASRPVPASRGKTLCKGDRQAPPGPARPPRPIMWSAPPRAPRSSTPCGGA 398
Qy      421  VREDTYPVGTGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVGEPMFRHYDFRQSCWV 480
Db      399  VREDTYPVGTGVSPALAQGGPQGSWRFLQWNSMPRLPTDLVGEPMFRHYDFRQSCWV 458
Qy      481  RAISOEDQLAPCQWQAEHPARVRSAPFAAPSTDSDQGTFFRARDPOCAPTSGPCLCGHL 540
Db      459  RAISOEDQLAPCQWQAEHPARVRSAPFAAPSTDSDQGTFFRARDPOCAPTSGPCLCGHL 518
Qy      541  ESSQPPPGF 549
Db      519  ESSQPPPGF 527

```

RESULT 11

ADM04434 ID ADM04434 standard; protein; 527 AA.

AC ADM04434;

DT 20-MAY-2004 (first entry)

DE Human protein of the invention SEQ ID NO:3119.

KM human; gene therapy; diagnostic marker; pharmaceutical.

OS Homo sapiens.

PN EP1347046-A1.

PD 24-SEP-2003.

PF 12-APR-2002; 2002EP-00008400.

PR 22-MAR-2002; 2002JP-00137785.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
 DR WPI; 2003-723558/69.

DR N-PSDB; ADM01991.
 XX New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX
 PS Claim 1; SEQ ID NO 3119; 3055bp; English.
 XX
 CC The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotide ADM03136-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC protein sequence of the invention.
 XX
 SQ Sequence 527 AA;
 Query Match 94.8%; Score 2827; DB 7; Length 527;
 Best Local Similarity 95.6%; Pred. No. 1.2e-243;
 Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;
 QY 1 MDVVVAGSWMAQREDDIMKYEKGHRAGLPEDKGPKEPRSYNNNVNDHGIHETELPPL 60
 DB 1 MDVVVAGSWMAQREDDIMKYEKGHRAGLPEDKGPKEPRSYNNNVNDHGIHETELPPL 60
 QY 61 TAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAVKGPMNIRGPMVSLNIEEM 120
 DB 61 TAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAVKGPMNIRGPMVSLNIEEM 120
 QY 121 KIKNGRQVQIMKEKGRSSEHIQRIDRVSGTLRKHIFFRDYRGYKORELHILLAYERY 180
 DB 121 KIKNGRQVQIMKEKGRSSEHIQRIDRVSGTLRKHIFFRDYRGYKORELHILLAYERY 180
 QY 121 KIKNGRQVQIMKEKGRSSEHIQRIDRVSGTLRKHIFFRDYRGYKORELHILLAYERY 180
 DB 121 KIKNGRQVQIMKEKGRSSEHIQRIDRVSGTLRKHIFFRDYRGYKORELHILLAYERY 180
 QY 181 NPEVGYCDLSHIALPLILYLPEDAFMVLVOLLSERHSLGFSFSPNGTVOGLDPOE 240
 DB 181 NPEVGYCDLSHIALPLILYLPEDAFMVLVOLLSERHSLGFSFSPNGTVOGLDPOE 240
 QY 241 HVVATSOPTKMGHODKDLGCGCSPGLGILRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
 DB 241 HVVATSOPTKMGHODKDLGCGCSPGLGILRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
 QY 301 TRIAFKVOQKRLITKTSRCGPMARFCNRPVDTWABDETVLKLRLASMKKLTROGDLPPP 360
 DB 279 TRIAFKVOQKRLITKTSRCGPMARFCNRPVDTWABDETVLKLRLASMKKLTROGDLPPP 338
 QY 361 AKPEGSSASRPVPSRSGKTLCKGDRQAPRPPAPRPPIWSASBPRAPRSTCPGGA 420
 DB 339 AKPEGSSASRPVPSRSGKTLCKGDRQAPRPPAPRPPIWSASBPRAPRSTCPGGA 398
 QY 421 VREDTVPVGTQGVPPAPALAQGGPQGSWRFQWNSMRLPTDLDVSGPMFRHDFQSCMV 480
 DB 399 VREDTVPVGTQGVPPAPALAQGGPQGSWRFQWNSMRLPTDLDVSGPMFRHDFQSCMV 458
 QY 481 RAISOEDQALPCWQAEHPAERVRSAFAPABSTDSOGTTPRANDEPCAPTSGPLCGLHL 540
 DB 459 RAISOEDQALPCWQAEHPAERVRSAFAPABSTDSOGTTPRANDEPCAPTSGPLCGLHL 518
 QY 541 ESSQPPRGP 549
 DB 519 ESSQPPRGP 527
 RESULT 12
 ABG97507 standard; protein; 549 AA.
 XX
 AC ABG97507;
 XX
 DT 16-DEC-2002 (first entry)

XX DE Human NOVX26 protein.
 XX
 KW Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
 KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
 KW tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
 KW Huntington's disease; cerebral palsy; epilepsy; Leach-Nyhan syndrome;
 KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
 KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
 KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
 KW scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
 KW asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
 KW bacterial infection; parasitic infection; graft-versus-host disease;
 KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
 KW angiogenesis.
 XX
 OS Homo sapiens.
 XX
 PN WO200272770-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US007283.
 XX
 PR 08-MAR-2001; 2001US-0274281P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275799P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0278152P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-MAY-2001; 2001US-0288148P.
 PR 31-MAY-2001; 2001US-0294821P.
 PR 31-OCT-2001; 2001US-0335302P.
 PR 04-DEC-2001; 2001US-0338375P.
 PR 07-MAR-2002; 2002US-00094466.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Spyrek KA, Vernet CA, Tchervet VT, Malyankar UM, Gerlach VL;
 PI Li U, Zernusen BD, Paturajan M, Gusev VY, Kekuda R, Pena CBA;
 PI Zhong M, Gangoli EA, Taupier RJ;
 XX
 DR WPI; 2002-113508/77.
 DR N-PSDB; AB878751.
 XX
 PT New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
 PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
 PT Parkinson's disease.
 XX
 PS Claim 1; Page 164; 266pp; English.
 XX
 CC The present invention relates to a new polypeptide (NOVX). The NOVX
 CC polypeptide, nucleic acid and antibody are useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease,
 CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
 CC polypeptides and antibodies are useful for treating, preventing or
 CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 CC palsy, epilepsy, Leach-Nyhan syndrome, multiple sclerosis, ataxia-
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,

CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
CC scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
CC infections, or graft-versus-host disease. The nucleic acids and
CC polypeptides may also be used as targets for the identification of small
CC molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, hematopoiesis, wound healing and
CC angiogenesis, in gene therapy, in generation of antibodies that bind
CC immunospecifically to NOVX substances for use in therapeutic or
CC diagnostic methods. The nucleic acids are further used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine, and
CC protein of the invention
CC
XX
SQ Sequence 549 AA;

Query Match 93.3%; Score 2784; DB 5; Length 549;
Best Local Similarity 93.8%; Pred. No. 8.7e-240;
Matches 515; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMAQERDIIIMKYEKGHRAGLPEDKPKPFRSYNNNVHDIGIVHETELPPL 60
DB 1 MDVVEVAGSWMAQERDIIIMKYEKGHRAGLPEDKPKPFRSYNNNVHDIGIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSWMDLGDWEKYSRKLIDRAYKGMNINRGPMNSVILNIEEM 120
DB 61 TAREAKQIRREISRSKSWMDLGDWEKYSRKLIDRAYKGMNINRGPMNSVILNIEEM 120
QY 121 KLNKPGRYOIMKEKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTOKORELHILHILAAVEEY 180
DB 121 KLNKPGRYOIMKEKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTOKORELHILHILAAVEEY 180
QY 181 NPEVGYCDLSHIALFLYLYPEEDAFWALVOLASRHSIQGPHSPNGGTVOGLDQOE 240
DB 181 NPEVGYCDLSHIALFLYLYPEEDAFWALVOLASRHSIQGPHSPNGGTVOGLDQOE 240
QY 241 HVAATSOPKTMGHODKDLGCGQSPGLGRLILIDGSLGLTLRLMDVYLVEBEOQLMPT 300
DB 241 HVAATSOPKTMGHODKDLGCGQSPGLGRLILIDGSLGLTLRLMDVYLVEBEOQLMPT 300
QY 301 TRIAFKVOQKRLTTSRCGPMARFCNRFVDTWADBDVTYLKHLRASMKLTKRQGLPP 360
DB 301 TRIAFKVOQKRLTTSRCGPMARFCNRFVDTWADBDVTYLKHLRASMKLTKRQGLPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPAPFPPIWSASPPRPSSTPCPGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPAPFPPIWSASPPRPSSTPCPGA 420
QY 421 VREDTYVGTQGVSPALAQGGQSWRFLQNNMSMRLLPTDLVDEGPMFRHYDFQSCMV 480
DB 421 VREDTYVGTQGVSPALAQGGQSWRFLQNNMSMRLLPTDLVDEGPMFRHYDFQSCMV 480
QY 481 RAISOEDOLAPCWAQHPARVRSAPAPASTSDGTPPRADBDPCAPTSGPCGLGHL 540
DB 481 RAISOEDOLAPCWAQHPARVRSAPAPASTSDGTPPRADBDPCAPTSGPCGLGHL 540
QY 541 ESSQPPPGF 549
DB 541 ESSQPPPGF 549

RESULT 13
ID ABG70738 standard; protein; 515 AA.

XX ABG70738;

XX 28-NOV-2002 (first entry)

DE Human PRC17 protein splice variant 2.

KW Human; PRC17; prostate cancer; ovarian cancer; lung cancer;

KW breast cancer; cytostatic; chromosome 17q11-12; splice variant.

XX Homo sapiens.

XX WO200262958-A2.

XX 15-AUG-2002.

XX 08-FEB-2002; 2002WO-US003457.

XX 08-FEB-2001; 2001US-0267615P.

XX (TULA-) TULARIK INC.

XX Li J, Powers S, Xiang P, Peng Y;

XX WPI; 2002-706902/76.

XX N-PSDB; ABS54708.

PT Novel isolated PRC17 polypeptide useful diagnostically or prognostically
PT to detect diseases or conditions associated with altered PRC17 activity
PT or expression relative to normal, for example cancer.

Claim 2; Page 65; 78pp; English.

CC The present invention relates to a new PRC17 polypeptide. The invention
CC is useful for detecting cancer cells (such as prostate tissue, breast
CC tissue, lung tissue, ovarian tissue) in a biological sample. The
CC invention is further useful for monitoring the efficacy of a therapeutic
CC treatment of a cancer (prostate cancer, ovarian cancer, lung cancer,
CC breast cancer). The molecules of the invention are useful diagnostically
CC or prognostically to detect diseases or conditions associated with
CC altered PRC17 activity or expression relative to normal, for example
CC cancer. The present amino acid sequence represents the human PRC17
CC protein splice variant 2 of the invention. The human PRC17 gene is
CC located on chromosome 17q11-12
CC
XX
SQ Sequence 515 AA;

Query Match 93.0%; Score 2773; DB 5; Length 515;
Best Local Similarity 93.8%; Pred. No. 7.7e-239;
Matches 515; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 MDVVEVAGSWMAQERDIIIMKYEKGHRAGLPEDKPKPFRSYNNNVHDIGIVHETELPPL 60
DB 1 MDVVEVAGSWMAQERDIIIMKYEKGHRAGLPEDKPKPFRSYNNNVHDIGIVHETELPPL 60
QY 61 TAREAKQIRREISRSKSWMDLGDWEKYSRKLIDRAYKGMNINRGPMNSVILNIEEM 120
DB 61 TAREAKQIRREISRSKSWMDLGDWEKYSRKLIDRAYKGMNINRGPMNSVILNIEEM 120
QY 121 KLNKPGRYOIMKEKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTOKORELHILHILAAVEEY 180
DB 121 KLNKPGRYOIMKEKRSSEHIQRIIDRVSGTLRKHIFFRDRYGTOKORELHILHILAAVEEY 180
QY 181 NPEVGYCDLSHIALFLYLYPEEDAFWALVOLASRHSIQGPHSPNGGTVOGLDQOE 240
DB 181 NPEVGYCDLSHIALFLYLYPEEDAFWALVOLASRHSIQGPHSPNGGTVOGLDQOE 240
QY 241 HVAATSOPKTMGHODKDLGCGQSPGLGRLILIDGSLGLTLRLMDVYLVEBEOQLMPT 300
DB 241 HVAATSOPKTMGHODKDLGCGQSPGLGRLILIDGSLGLTLRLMDVYLVEBEOQLMPT 300
QY 301 TRIAFKVOQKRLTTSRCGPMARFCNRFVDTWADBDVTYLKHLRASMKLTKRQGLPP 360
DB 301 TRIAFKVOQKRLTTSRCGPMARFCNRFVDTWADBDVTYLKHLRASMKLTKRQGLPP 360
QY 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPAPFPPIWSASPPRPSSTPCPGA 420
DB 361 AKPEQSSASRPVPASRGKTLCKGDRQAPPGPAPFPPIWSASPPRPSSTPCPGA 420
QY 421 VREDTYVGTQGVSPALAQGGQSWRFLQNNMSMRLLPTDLVDEGPMFRHYDFQSCMV 480
DB 421 VREDTYVGTQGVSPALAQGGQSWRFLQNNMSMRLLPTDLVDEGPMFRHYDFQSCMV 480

```
Db 387 VREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMRLPTDLDEGPMFRHYDFRQSCMV 446
Qy 481 RAISOEDOLAPCQMAEHPAERVSAPAPASTDSOGTPRAREDEPCATSPGCLGLH 540
Db 447 RAISOEDOLAPCQMAEHPAERVSAPAPASTDSOGTPRAREDEPCATSPGCLGLH 506
Qy 541 ESSQFPFPGF 549
Db 507 ESSQFPFPGF 515

RESULT 14
ABUS3235
ID ABUS3235 standard; protein; 499 AA.
XX AC ABUS3235;
XX DT 14-APR-2003 (first entry)
XX DE Human cell cycle-associated DKFZphes3_35g22 homologue.
XX KM Human; gene therapy; vaccine; disease treatment; detection.
XX OS Homo sapiens.
XX PN WO200112659-A2.
XX PD 22-FEB-2001.
XX PF 18-AUG-2000; 2000WO-1B001496.
XX PR 18-AUG-1999; 99US-0149499P.
XX PR 28-SEP-1999; 99US-0156503P.
XX PA (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX PI Wiemann S;
XX DR MPI; 2001-327840/34.
XX PT Nucleic acids having the sequences of clones isolated from libraries of
XX PT different human tissues, useful in recombinant DNA methodologies.
XX PS Example III; Page 868-869; 1095pp; English.
XX CC This invention describes novel polynucleotides and polypeptides isolated
XX CC from human cDNA libraries which can be used for gene therapy or in
XX CC vaccines. The polynucleotides of the invention and antibodies encoded by
XX CC them may be used in the prevention, diagnosis and treatment of diseases
XX CC associated with inappropriate polypeptide expression. The products of the
XX CC invention may also be used to identify modulators of expression and
XX CC activity and to down regulate expression and activity. The antibodies of
XX CC the invention may also be used as diagnostic agents for detecting the
XX CC presence of polypeptides in samples. This sequence represents a homologue
XX CC of a polypeptide described in the disclosure of the invention
XX SQ Sequence 499 AA;

Query Match 73.5%; Score 2191.5; DB 4; Length 499;
Best Local Similarity 81.6%; Pred. No. 7.6e-187;
Matches 408; Conservative 33; Mismatches 58; Indels 1; Gaps 1;

Qy 1 MOVVEVAGSWMAQREDITMKTEKGRALPEDKGPSPRSTNNVNDHGIYHETLPL 60
Db 1 MDMVENAOSLQAOERDILMKYDKGRAGLPEDKGPSPV-GINSSIDRGILHETLPPV 59
Qy 61 TAREAKOIRREISRSKSWTMDLGMKEYSKRLIDRAYKGMPTNRGPMWVLTALNIBM 120
Db 60 TAREAKKIRREITRSKSWTMDLGMKEYSKRLIDRAYKGMPTNRGPMWVLTALNIBM 119
Qy 121 KLNKNGRYQIMKGRKSSSEHIQRIDRVSGTLRKAIIPFRDRYGTQKRELLHILAYEY 180
Db 120 KLNKNGRYQIMKGRKSSSEHIQRIDRVSGTLRKAIIPFRDRYGTQKRELLHILAYEY 179
```

```
Qy 181 NPEVGYCRDLSHIALFLYLPEEDAFMALVOLLASERHSLQGFSPNGATVQGLDQOE 240
Db 180 NPEVGYCRDLSHITLFLYLPEEDAFMALVOLLASERHSLQGFSPNGATVQGLDQOE 239
Qy 241 HVAATSOPEKTMGHODKDLQGGCSPGLIRLIDGISLGLTLRLMDVYLVEGQALMPI 300
Db 240 HVAATSOPEKTMGHODKDLQGGCSPGLIRLIDGISLGLTLRLMDVYLVEGQALMPI 299
Qy 301 TRIAPRVQOKRLTKTRSGCPMFAFCRFPVDTARBEDTYLKLRLASMKULTKQGLDPP 360
Db 300 TSIALKVQOKRLTKTRSGCPMFAFCRFPVDTARBEDTYLKLRLASMKULTKQGLDPP 359
Qy 361 AKPEQSSASRPVPSPRGGKTLCKGDRAPCPGPAPRPPIWASPPRARSSTPCPGA 420
Db 360 AKPEQSSASRPVPSPRGGKTLCKGDRAPCPGPAPRPPIWASPPRARSSTPCPGA 419
Qy 421 VREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMRLPTDLDEGPMFRHYDFRQSCMV 480
Db 420 VREDTYPVGTGVPSPALAQGGPQGSWRFLOWNSMRLPTDLDEGPMFRHYDFRQSCMV 479
Qy 481 RAISOEDOLAPCQMAEHPA 500
Db 480 RAISOEDOLAPCQMAEHPA 499

RESULT 15
ABG24026
ID ABG24026 standard; protein; 805 AA.
XX AC ABG24026;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #24017.
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KM food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HSE-) HSE INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR N-PSDB; AAS88213.
XX PT MPI; 2001-639362/73.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 54385; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (II) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (II) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
```

CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostic, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG3037 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 805 AA;

Query Match 73.5%; Score 2191.5; DB 4; Length 805;
 Best Local Similarity 81.6%; Pred. No. 1.5e-186;
 Matches 408; Conservative 33; Mismatches 58; Indels 1; Gaps 1;

QY 1 MDVVEVAGSWMAOEREDITMKYEGHRAGLPEDKGPKEPRSYNNNVNDHGIYHETELPPL 60
 DB 20 MDVVEVAGSWMAOEREDITMKYEGHRAGLPEDKGPKEPRSYNNNVNDHGIYHETELPPL 78
 QY 61 TAREAKOIRREISRSKSKWVDMGDWEKYSRRLIDRAYKGMNIRGPMGVLNIBEM 120
 DB 79 TAREAKOIRREISRSKSKWVDMGDWEKYSRRLIDRAYKGMNIRGPMGVLNIBEM 138
 QY 121 KLKNGRGYQIMKEKGRSESEH1QRIDRVSGTLRKHIFFPRDRYGTQORELHILLAYEY 180
 DB 139 KLKNGRGYQIMKEKGRSESEH1IHIDLVRTLRNHFPRDRYGAQORELFYILLAYSEY 198
 QY 181 NPEVGYCRPLSHIAFLIYLPEDEAFALVOLLASERHSLOGFHSPNGTVOGLDQOE 240
 DB 199 NPEVGYCRPLSHIAFLIYLPEDEAFALVOLLASERHSLOGFHSPNGTVOGLDQOE 258
 QY 241 HVVATSQPTMGHODKDLGQGSPLGCTIRILIDGISLGLTLRLMDVYLVEGEQALMPI 300
 DB 259 HVVATSQPTMGHODKDLGQGSPLGCTIRILIDGISLGLTLRLMDVYLVEGEQALMPI 318
 QY 301 TRIAFVQQRILTKTSRCGPWAFENRFVDTWARDEDTVLKHLRASMKULTRKQDLP 360
 DB 319 TRIAFVQQRILTKTSRCGPWAFENRFVDTWARDEDTVLKHLRASMKULTRKQDLP 378
 QY 361 AKPEQSSASRPVAPARGGKTLCKGRDRAPEGPARFPRPIWASPPRAPRSSTPCPGA 420
 DB 379 AKPEQSSASRPVAPARGGKTLCKGRDRAPEGPARFPRPIWASPPRAPRSSTPCPGA 438
 QY 421 VREDTYPVGTVGVPSPALAQGPGQSWRFLQWNSMPRLPTDLDVEGPFWRHYDFRQSCWV 480
 DB 439 VREDTYPVGTVGVPSPALAQGPGQSWRFLQWNSMPRLPTDLDIGGPFWRHYDFRQSCWV 498
 QY 481 RAISQEDQLAPCWAEPAPAE 500
 DB 499 RAISQEDQLAPCWAEPAPAE 518

Search completed: February 4, 2005, 05:08:18
 Job time : 92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2005, 05:10:13 ; Search time 84 Seconds
(without alignments)
2361.284 Million cell updates/sec

Title: US-10-071-838-2

Perfect score: 2983
Sequence: 1 MDVVEVAGSMWAQEREDIM.....TSGPCLGHLHSSQFPFPGF 549

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2983	100.0	549	US-10-071-838-2	Sequence 2, Appl1
2	2978	99.8	549	US-10-839-882-12	Sequence 12, Appl1
3	2952	99.0	549	US-10-071-838-18	Sequence 18, Appl1
4	2934.5	98.4	610	US-10-071-838-4	Sequence 4, Appl1
5	2827	94.3	527	US-10-108-2604-3119	Sequence 3119, Ap
6	2784	93.8	549	US-10-094-466-62	Sequence 62, Appl1
7	2773	93.0	515	US-10-071-838-6	Sequence 6, Appl1
8	2191.5	79.5	786	US-10-071-838-17	Sequence 17, Appl1
9	433.5	14.5	210	US-10-108-2604-4446	Sequence 4446, Ap
10	348.5	11.7	381	US-10-425-115-237632	Sequence 237632, A
11	348.5	11.7	434	US-10-425-114-64834	Sequence 64834, A
12	348.5	11.7	434	US-10-425-114-70559	Sequence 70559, A
13	329	11.0	433	US-10-437-963-139162	Sequence 139162, A

14	321.5	10.8	537	15	US-10-051-874-138	Sequence 138, App
15	320.5	10.7	533	15	US-10-051-874-137	Sequence 137, App
16	320.5	10.7	533	15	US-10-408-765A-1052	Sequence 1052, Ap
17	305.5	10.2	500	15	US-10-051-874-135	Sequence 135, App
18	299.5	10.0	508	15	US-10-051-874-46	Sequence 46, Appl1
19	299.5	10.0	508	15	US-10-051-874-134	Sequence 134, App
20	298	10.0	438	15	US-10-051-874-136	Sequence 136, App
21	290.5	9.7	241	16	US-10-767-701-32143	Sequence 32143, A
22	282	9.5	289	15	US-10-424-599-148597	Sequence 148597, A
23	266.5	8.9	209	15	US-10-051-874-139	Sequence 139, App
24	231	7.7	44	9	US-09-864-761-43564	Sequence 43564, A
25	229.5	7.7	811	16	US-10-437-963-118949	Sequence 118949, A
26	226.5	7.6	723	16	US-10-322-281-455	Sequence 455, App
27	226	7.6	806	16	US-10-322-281-458	Sequence 458, App
28	222	7.4	607	17	US-10-370-715B-328	Sequence 328, App
29	214	7.2	468	14	US-10-258-883-2	Sequence 2, Appl1
30	214	7.2	924	14	US-10-177-293-244	Sequence 244, App
31	214	7.2	928	11	US-09-989-890-198	Sequence 198, App
32	214	7.2	928	16	US-10-473-576-16	Sequence 16, Appl1
33	214	7.2	1004	14	US-10-094-466-58	Sequence 58, Appl1
34	214	7.2	1026	16	US-10-467-595-5	Sequence 5, Appl1
35	209	7.0	768	9	US-09-764-868-728	Sequence 728, App
36	209	7.0	768	11	US-09-764-875-922	Sequence 922, App
37	202.5	6.8	815	15	US-10-108-260A-3374	Sequence 3374, App
38	201.5	6.8	775	15	US-10-112-944-291	Sequence 291, App
39	201.5	6.8	775	15	US-10-112-944-749	Sequence 749, App
40	200	6.7	917	14	US-10-221-658-4	Sequence 4, Appl1
41	199.5	6.7	814	15	US-10-158-057-243	Sequence 243, App
42	198	6.6	756	11	US-09-764-875-847	Sequence 847, App
43	197.5	6.6	366	15	US-10-264-237-1756	Sequence 1756, Ap
44	197.5	6.6	373	9	US-09-764-848-32	Sequence 32, Appl1
45	197.5	6.6	373	14	US-10-116-016-32	Sequence 32, Appl1

ALIGNMENTS

RESULT 1
US-10-071-838-2
Sequence 2, Application US/10071838
Publication No. US20030044814A1
GENERAL INFORMATION:
APPLICANT: Li, Jing
APPLICANT: Powers, Scott
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
FILE REFERENCE: 018781-007610US
CURRENT APPLICATION NUMBER: US/10/071,838
CURRENT FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: US 60/267,615
PRIOR FILING DATE: 2001-02-08
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 549
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human PRC17
US-10-071-838-2
Query Match 100.0%; Score 2983; DB 14; Length 549;
Best Local Similarity 100.0%; Pred. No. 6.4e-225;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDVVEVAGSMWAQEREDIMKYCKGRAGLPEDCKPKFRSYNNVVDHIGYHETELPPL 60
DB 1 MDVVEVAGSMWAQEREDIMKYCKGRAGLPEDCKPKFRSYNNVVDHIGYHETELPPL 60
QY TAREAKQIRREISRSKXVVDMLGDWEKXSSRKLIDRAKYKMPNIRGPMNSVLINIEEM 120

Db 61 TAREAQIRREISRSKSWMDLGMKYSRKLIDRAVKGMPNIRGPMWSVLLNTEEM 120
Qy 121 KLNKRGYQIMKEKGRSSSEHIORIDRVSGTLRKHIIPRDYGTQRELIHLLAYEY 180
Db 121 KLNKRGYQIMKEKGRSSSEHIORIDRVSGTLRKHIIPRDYGTQRELIHLLAYEY 180
Qy 181 NPEVGCRLSHIALFLYLPEEDAFMALVOLLAESRHSLOGFHSFNGCTVQGLQDOE 240
Db 181 NPEVGCRLSHIALFLYLPEEDAFMALVOLLAESRHSLOGFHSFNGCTVQGLQDOE 240
Qy 241 HVVATSQPTMGHODKDLCCGQCSPLGCLIRLLIDGISTGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPTMGHODKDLCCGQCSPLGCLIRLLIDGISTGLTLRLMDVYLVEGEQALMPI 300
Qy 301 TRIAFYQOKRLTKTSRCGPWFARFCNRFVDTWABEDTVLKHLSMSKLTTRKQGLPPP 360
Db 301 TRIAFYQOKRLTKTSRCGPWFARFCNRFVDTWABEDTVLKHLSMSKLTTRKQGLPPP 360
Qy 361 AKPEQSSASRPVPSRGGKTLCKGDRQAPGPPARFPRIWASPPRARSSTPCPGA 420
Db 361 AKPEQSSASRPVPSRGGKTLCKGDRQAPGPPARFPRIWASPPRARSSTPCPGA 420
Qy 421 VREDTYPVGTQVSPALAOGGPQSWRFLQWNSMPLPTDLVGEPMFRHYDFRQSCW 480
Db 421 VREDTYPVGTQVSPALAOGGPQSWRFLQWNSMPLPTDLVGEPMFRHYDFRQSCW 480
Qy 481 RAISOEDOLAPCWAHEPARVRSAPAPSTDSOGTFPRARBOCAPTSGBCLCGHL 540
Db 481 RAISOEDOLAPCWAHEPARVRSAPAPSTDSOGTFPRARBOCAPTSGBCLCGHL 540
Qy 541 ESSQPPPGF 549
Db 541 ESSQPPPGF 549

RESULT 2

US-10-839-882-12
; Sequence 12, Application US/10839882
; Publication No. US20040203106A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: TANG, Y. Tom
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: GUEGLER, Karl J.
; APPLICANT: CORLEY, Neil C.
; APPLICANT: LAL, Preeti
; APPLICANT: AZIMZAI, Yajda
; APPLICANT: BRUGH, Mariah R.
; APPLICANT: JUNG, Yang
; APPLICANT: SHIH, Leo L.
; TITLE OF INVENTION: PROLIFERATION AND APOPTOSIS RELATED PROTEINS
; FILE REFERENCE: PF-0619 PCT
; CURRENT APPLICATION NUMBER: US/10/839,882
; PRIOR FILING DATE: 2004-05-05
; PRIOR APPLICATION NUMBER: US/09/807,452
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/175,737, unaassigned; 60/118,559; 09/249,740, unaassigned;
; 60/154,336
; PRIOR FILING DATE: 1998-10-20; 1998-10-20; 1999-02-04; 1999-04-11; 1999-04-11;
; 1999-04-22
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4325626CD1
US-10-839-882-12

Query Match 99.8%; Score 2978; DB 17; Length 549;

Confusion

Best Local Similarity 99.8%; Pred. No. 1,6e-224;
Matches 548; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MDVVEVAGSWMAOERDIIIMKEKGRAGLPEDKGPFRSNNVNDHGYHETELPPL 60
Db 1 MDVVEVAGSWMAOERDIIIMKEKGRAGLPEDKGPFRSNNVNDHGYHETELPPL 60
Qy 61 TAREAQIRREISRSKSWMDLGMKYSRKLIDRAVKGMPNIRGPMWSVLLNTEEM 120
Db 61 TAREAQIRREISRSKSWMDLGMKYSRKLIDRAVKGMPNIRGPMWSVLLNTEEM 120
Qy 121 KLNKRGYQIMKEKGRSSSEHIORIDRVSGTLRKHIIPRDYGTQRELIHLLAYEY 180
Db 121 KLNKRGYQIMKEKGRSSSEHIORIDRVSGTLRKHIIPRDYGTQRELIHLLAYEY 180
Qy 181 NPEVGCRLSHIALFLYLPEEDAFMALVOLLAESRHSLOGFHSFNGCTVQGLQDOE 240
Db 181 NPEVGCRLSHIALFLYLPEEDAFMALVOLLAESRHSLOGFHSFNGCTVQGLQDOE 240
Qy 241 HVVATSQPTMGHODKDLCCGQCSPLGCLIRLLIDGISTGLTLRLMDVYLVEGEQALMPI 300
Db 241 HVVATSQPTMGHODKDLCCGQCSPLGCLIRLLIDGISTGLTLRLMDVYLVEGEQALMPI 300
Qy 301 TRIAFYQOKRLTKTSRCGPWFARFCNRFVDTWABEDTVLKHLSMSKLTTRKQGLPPP 360
Db 301 TRIAFYQOKRLTKTSRCGPWFARFCNRFVDTWABEDTVLKHLSMSKLTTRKQGLPPP 360
Qy 361 AKPEQSSASRPVPSRGGKTLCKGDRQAPGPPARFPRIWASPPRARSSTPCPGA 420
Db 361 AKPEQSSASRPVPSRGGKTLCKGDRQAPGPPARFPRIWASPPRARSSTPCPGA 420
Qy 421 VREDTYPVGTQVSPALAOGGPQSWRFLQWNSMPLPTDLVGEPMFRHYDFRQSCW 480
Db 421 VREDTYPVGTQVSPALAOGGPQSWRFLQWNSMPLPTDLVGEPMFRHYDFRQSCW 480
Qy 481 RAISOEDOLAPCWAHEPARVRSAPAPSTDSOGTFPRARBOCAPTSGBCLCGHL 540
Db 481 RAISOEDOLAPCWAHEPARVRSAPAPSTDSOGTFPRARBOCAPTSGBCLCGHL 540
Qy 541 ESSQPPPGF 549
Db 541 ESSQPPPGF 549

RESULT 3

US-10-071-838-18
; Sequence 18, Application US/10071838
; Publication No. US2003004481A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071,838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; PRIOR FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17.codi
US-10-071-838-18

Query Match 99.0%; Score 2952; DB 14; Length 549;
Best Local Similarity 99.1%; Pred. No. 1.7e-222;
Matches 544; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMAQEBEDIIIMKEKGRAGLPEDKGPCKPFRSYNNVNDHLGIHETELPPL 60
DB 1 MDVVEVAGSWMAQEBEDIIIMKEKGRAGLPEDKGPCKPFRSYNNVNDHLGIHETELPPL 60
QY 61 TAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
DB 61 TAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
QY 121 KLNKNGRQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQKORELHILLAYEEY 180
DB 121 KLNKNGRQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQKORELHILLAYEEY 180
QY 181 NPEVGYCRDLSHIALFLYLPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLQDOOQ 240
DB 181 NPEVGYCRDLSHIALFLYLPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLQDOOQ 240
QY 241 HNVATSOPTKMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOALMP 300
DB 241 HNVATSOPTKMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOALMP 300
QY 301 TRIAFKVOQKRLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLASMKKLTTRKQGLPP 360
DB 301 TRIAFKVOQKRLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLASMKKLTTRKQGLPP 360
QY 361 AKPEGSSASRPVPSRSGKTLCKGDRQAPGPPARPPRPISASPPRASPSTPCPGSA 420
DB 361 AKPEGSSASRPVPSRSGKTLCKGDRQAPGPPARPPRPISASPPRASPSTPCPGSA 420
QY 421 VREDTYPVGTGQVSPALAOGPGSWMPLQWNSMPLPTDLVGEPMFRHYDPROSCW 480
DB 421 VREDTYPVGTGQVSPALAOGPGSWMPLQWNSMPLPTDLVGEPMFRHYDPROSCW 480
QY 481 RAISOEDOLAPCWAHEPAERVSAPAPSTDSOGTFFRANDEPCAPTSGPCLCGHL 540
DB 481 RAISOEDOLAPCWAHEPAERVSAPAPSTDSOGTFFRANDEPCAPTSGPCLCGHL 540
QY 541 ESSQPPPGF 549
DB 541 ESSQPPPGF 549

RESULT 4
US-10-071-838-4
; Sequence 4, Application US/10071838
; Publication No. US2003004814A1
; GENERAL INFORMATION:
; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Kiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071, 838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human PRC17 splice variant 1 (exon 3 expanded)
US-10-071-838-4

Query Match 98.4%; Score 2934.5; DB 14; Length 610;
Best Local Similarity 89.8%; Pred. No. 4, 6e-221;
Matches 548; Conservative 0; Mismatches 1; Indels 61; Gaps 1;
QY 1 MDVVEVAGSWMAQEBEDIIIMKEKGRAGLPEDKGPCKPFRSYNNVNDHLGIH----- 53

DB 1 MDVVEVAGSWMAQEBEDIIIMKEKGRAGLPEDKGPCKPFRSYNNVNDHLGIHOSCRWS 60
QY 54 ----- -TELEPP 59
DB 61 APQECPPFPVPSDGLPELERDRASPFWGAPLGLQAPCSSALPGLYSTELPP 120
QY 60 LTAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLLNIEE 119
DB 121 LTAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLLNIEE 180
QY 120 KLNKNGRQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQKORELHILLAYEE 179
DB 181 KLNKNGRQIMKEKGRSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQKORELHILLAYEE 240
QY 180 NPEVGYCRDLSHIALFLYLPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLQDOOQ 239
DB 241 NPEVGYCRDLSHIALFLYLPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLQDOOQ 300
QY 240 HNVATSOPTKMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOALMP 299
DB 301 HNVATSOPTKMGHODKDLGQCSPLGCLIRILIDGISLGLTLRLMDVYLVEGEOALMP 360
QY 300 TRIAFKVOQKRLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLASMKKLTTRKQGLPP 359
DB 361 TRIAFKVOQKRLTKTSRCGPMARFCNRFVDTWARDEDTVLKHLASMKKLTTRKQGLPP 420
QY 360 PAKPEGSSASRPVPSRSGKTLCKGDRQAPGPPARPPRPISASPPRASPSTPCPGSA 419
DB 421 PAKPEGSSASRPVPSRSGKTLCKGDRQAPGPPARPPRPISASPPRASPSTPCPGSA 480
QY 420 AVREDTYPVGTGQVSPALAOGPGSWMPLQWNSMPLPTDLVGEPMFRHYDPROSCW 479
DB 481 AVREDTYPVGTGQVSPALAOGPGSWMPLQWNSMPLPTDLVGEPMFRHYDPROSCW 540
QY 480 VRAISOEDOLAPCWAHEPAERVSAPAPSTDSOGTFFRANDEPCAPTSGPCLCGHL 539
DB 541 VRAISOEDOLAPCWAHEPAERVSAPAPSTDSOGTFFRANDEPCAPTSGPCLCGHL 600
QY 540 LESSQPPPGF 549
DB 601 LESSQPPPGF 610

RESULT 5
US-10-108-260A-3119
; Sequence 3119, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3119
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3119

Query Match 94.8%; Score 2827; DB 15; Length 527;
Best Local Similarity 95.6%; Pred. No. 1e-212;
Matches 525; Conservative 0; Mismatches 2; Indels 22; Gaps 1;
QY 1 MDVVEVAGSWMAQEBEDIIIMKEKGRAGLPEDKGPCKPFRSYNNVNDHLGIHETELPPL 60
DB 1 MDVVEVAGSWMAQEBEDIIIMKEKGRAGLPEDKGPCKPFRSYNNVNDHLGIHETELPPL 60
QY 61 TAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLLNIEEM 120
DB 61 TAREAKQIRREISRSKSNVMDLGDWEKYSRKLIDRAYKGMPMNIRGPMWSVLLNIEEM 120

QY 121 KLNKRGYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQRELIHILLAYEY 180
DB 121 KLNKRGYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQRELIHILLAYEY 180
QY 181 NPEVGYCRDLISHIALFLYLYPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLDOOE 240
DB 181 NPEVGYCRDLISHIALFLYLYPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLDOOE 240
QY 241 HVATSOPTKMGHODKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEOLMP1 300
DB 241 HVATSOPTKMGHODKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEOLMP1 300
QY 301 TRIAFVQOQRLTKTSRCGPWFARFCNRFVDTWARDEDTYLKHLRASMKKLTRKQGLPPP 360
DB 279 TRIAFVQOQRLTKTSRCGPWFARFCNRFVDTWARDEDTYLKHLRASMKKLTRKQGLPPP 360
QY 361 AKPEQSSASRPVPSASRGKTLCKGDRQAPPGPPARPPRPIWSASPPRAPSSTPCPGA 420
DB 339 AKPEQSSASRPVPSASRGKTLCKGDRQAPPGPPARPPRPIWSASPPRAPSSTPCPGA 420
QY 421 VREDTYPVGTQVPSPALAQGPGSWRFLQWNSMRLPTDLVGEPMFRHYDFRQSCWV 480
DB 399 VREDTYPVGTQVPSPALAQGPGSWRFLQWNSMRLPTDLVGEPMFRHYDFRQSCWV 480
QY 481 RAISQEDQAPCQOAEHPAERVRSAPASTSDQGTFFRARDQPCAPTSGPCLCGLHL 540
DB 459 RAISQEDQAPCQOAEHPAERVRSAPASTSDQGTFFRARDQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPFPGF 549
DB 519 ESSQFPFPGF 527

RESULT 6

US-10-094-466-62
Sequence 62, Application US/10094466
Publication No. US2003020363A1
GENERAL INFORMATION:
APPLICANT: Spytek et al.
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
TITLE OF INVENTION: AND METHODS OF USING
FILE REFERENCE: 21402-290D
CURRENT APPLICATION NUMBER: US/10/094,466
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/288,148
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/338,375
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 60/275,579
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/335,302
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 60/275,601
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/276,000
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/277,338
PRIOR FILING DATE: 2001-03-20
Prior Application data removed - See file Wrapper or PAM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patin 2.1
SEQ ID NO 62
LENGTH: 549
TYPE: PRT
ORGANISM: Homo sapiens

US-10-094-466-62

Query Match 93.3%; Score 2784; DB 14; Length 549;
Best Local Similarity 93.8%; Pred. No. 2,5e-209;
Matches 515; Conservative 13; Mismatches 21; Indels 0; Gaps 0;

QY 1 MDVVEVAGSWMAQERDITIMKYEKGHRAGLPEDKPKPPRSNNNVNHLGIYHETELPPL 60
DB 1 MDVVEVAGSWMAQERDITIMKYEKGHRAGLPEDKPKPPRSNNNVNHLGIYHETELPPL 60
QY 61 TAREVQOIRREISRSKSWYDMLGDWEKYSRKLIDRAYKGMNIRGPMNVSLLNIBEM 120
DB 61 TAREVQOIRREISRSKSWYDMLGDWEKYSRKLIDRAYKGMNIRGPMNVSLLNIBEM 120
QY 121 KLNKRGYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQRELIHILLAYEY 180
DB 121 KLNKRGYQIMKEKGRSSEHIQRIIDRVSGTLRKHIFFRDRYGTQRELIHILLAYEY 180
QY 181 NPEVGYCRDLISHIALFLYLYPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLDOOE 240
DB 181 NPEVGYCRDLISHIALFLYLYPEEDAFWALVOLLASERHSLOGFHSFNGGTVOGLDOOE 240
QY 241 HVATSOPTKMGHODKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEOLMP1 300
DB 241 HVATSOPTKMGHODKDLGQCSPLGCLIRILIDISLGLTLRLMDVYLVEGEOLMP1 300
QY 301 TRIAFVQOQRLTKTSRCGPWFARFCNRFVDTWARDEDTYLKHLRASMKKLTRKQGLPPP 360
DB 301 TRIAFVQOQRLTKTSRCGPWFARFCNRFVDTWARDEDTYLKHLRASMKKLTRKQGLPPP 360
QY 361 AKPEQSSASRPVPSASRGKTLCKGDRQAPPGPPARPPRPIWSASPPRAPSSTPCPGA 420
DB 361 AKPEQSSASRPVPSASRGKTLCKGDRQAPPGPPARPPRPIWSASPPRAPSSTPCPGA 420
QY 421 VREDTYPVGTQVPSPALAQGPGSWRFLQWNSMRLPTDLVGEPMFRHYDFRQSCWV 480
DB 421 VREDTYPVGTQVPSPALAQGPGSWRFLQWNSMRLPTDLVGEPMFRHYDFRQSCWV 480
QY 481 RAISQEDQAPCQOAEHPAERVRSAPASTSDQGTFFRARDQPCAPTSGPCLCGLHL 540
DB 481 RAISQEDQAPCQOAEHPAERVRSAPASTSDQGTFFRARDQPCAPTSGPCLCGLHL 540
QY 541 ESSQFPFPGF 549
DB 541 ESSQFPFPGF 549

RESULT 7

US-10-071-838-6
Sequence 6, Application US/10071838
Publication No. US2003004481A1
GENERAL INFORMATION:
APPLICANT: Li, Jing
APPLICANT: Powers, Scott
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: PRC17: An Amplified Cancer Gene
FILE REFERENCE: 018781-007610US
CURRENT APPLICATION NUMBER: US/10/071,838
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: US 60/267,615
PRIOR FILING DATE: 2001-02-08
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 515
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human PRC17 splice variant 2 (exon 10 deleted)
US-10-071-838-6

Query Match 93.0%; Score 2773; DB 14; Length 515;
Best Local Similarity 93.8%; Pred. No. 1.7e-208;
Matches 515; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

```
QY 1 MDVVEASWMAQOEEDIIIMKEKGRAGLPEDKGPCKPRSYNNNVHIGIYHETELPPL 60
DB 1 MDVVEASWMAQOEEDIIIMKEKGRAGLPEDKGPCKPRSYNNNVHIGIYHETELPPL 60
QY 61 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 120
QY 61 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 120
QY 121 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 180
DB 121 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 180
QY 121 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 180
DB 121 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 180
QY 181 NPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLOGFHSPNGTVQGLQDOE 240
DB 181 NPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLOGFHSPNGTVQGLQDOE 240
QY 241 HVVATSOPTMGHODKDLCCGCSPLGCLIRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSOPTMGHODKDLCCGCSPLGCLIRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 241 HVVATSOPTMGHODKDLCCGCSPLGCLIRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
DB 241 HVVATSOPTMGHODKDLCCGCSPLGCLIRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
QY 301 TRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDDETVLKLRLASMKKLTRKQGLDPPP 360
DB 301 TRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDDETVLKLRLASMKKLTRKQGLDPPP 360
QY 267 TRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDDETVLKLRLASMKKLTRKQGLDPPP 326
DB 267 TRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDDETVLKLRLASMKKLTRKQGLDPPP 326
QY 361 AKPEGSSASRPVPASRGSKTLCKGDRQAPGPPARFPRPIWSASPPAPRSSTPCPGA 420
DB 361 AKPEGSSASRPVPASRGSKTLCKGDRQAPGPPARFPRPIWSASPPAPRSSTPCPGA 420
QY 327 AKPEGSSASRPVPASRGSKTLCKGDRQAPGPPARFPRPIWSASPPAPRSSTPCPGA 386
DB 327 AKPEGSSASRPVPASRGSKTLCKGDRQAPGPPARFPRPIWSASPPAPRSSTPCPGA 386
QY 421 VREDTYPVGTGCVSPALAOGGPOGSMRFLQMSNMPRLPTDLVVGPMFRHDFRQSCMV 480
DB 421 VREDTYPVGTGCVSPALAOGGPOGSMRFLQMSNMPRLPTDLVVGPMFRHDFRQSCMV 480
QY 387 VREDTYPVGTGCVSPALAOGGPOGSMRFLQMSNMPRLPTDLVVGPMFRHDFRQSCMV 446
DB 387 VREDTYPVGTGCVSPALAOGGPOGSMRFLQMSNMPRLPTDLVVGPMFRHDFRQSCMV 446
QY 481 RAISOEDQALPCWQAEHPAEVRSAFAAPSTSDPGTFPRADDEPCAPTSGPCGLHL 540
DB 481 RAISOEDQALPCWQAEHPAEVRSAFAAPSTSDPGTFPRADDEPCAPTSGPCGLHL 540
QY 447 RAISOEDQALPCWQAEHPAEVRSAFAAPSTSDPGTFPRADDEPCAPTSGPCGLHL 506
DB 447 RAISOEDQALPCWQAEHPAEVRSAFAAPSTSDPGTFPRADDEPCAPTSGPCGLHL 506
QY 541 ESSQPPPGF 549
DB 507 ESSQPPPGF 515
```

RESULT 8
US-10-071-838-17
; Sequence 17, Application US/10071838
; Publication No. US20030044814A1
; GENERAL INFORMATION:

; APPLICANT: Li, Jing
; APPLICANT: Powers, Scott
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: PRC17. An Amplified Cancer Gene
; FILE REFERENCE: 018781-007610US
; CURRENT APPLICATION NUMBER: US/10/071.838
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/267,615
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 786
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: human TRE-2/USP6
US-10-071-838-17

Query Match 73.5%; Score 2191.5; DB 14; Length 786;
Best Local Similarity 81.6%; Pred. No. 1e-162;

Matches 408; Conservative 33; Mismatches 58; Indels 1; Gaps 1;

```
QY 1 MDVVEASWMAQOEEDIIIMKEKGRAGLPEDKGPCKPRSYNNNVHIGIYHETELPPL 60
DB 1 MDVVEASWMAQOEEDIIIMKEKGRAGLPEDKGPCKPRSYNNNVHIGIYHETELPPL 60
QY 61 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 120
DB 61 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 120
QY 60 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 119
DB 60 TAREAKQIRREISRSKSNVDMIGDWEKYSRKLIDRAYKGMPMNIRGPMVSVLNIEEM 119
QY 121 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 180
DB 121 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 179
QY 120 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 179
DB 120 KLKNGRQYQIMKEKGRSSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEY 179
QY 181 NPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLOGFHSPNGTVQGLQDOE 240
DB 180 NPEVGYCRDLISHIALFLLYLPEBDAFWALVOLLASERHSLOGFHSPNGTVQGLQDOE 239
QY 241 HVVATSOPTMGHODKDLCCGCSPLGCLIRLIDGISLGLTLRLMDVYLVEGEQALMPI 300
DB 240 HVVATSOPTMGHODKDLCCGCSPLGCLIRLIDGISLGLTLRLMDVYLVEGEQALMPI 299
QY 301 TRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDDETVLKLRLASMKKLTRKQGLDPPP 360
DB 300 TRIAFKVOQKRLTKTSRCGPWAFPCNRFVDTWARDDETVLKLRLASMKKLTRKQGLDPPP 359
QY 361 AKPEGSSASRPVPASRGSKTLCKGDRQAPGPPARFPRPIWSASPPAPRSSTPCPGA 420
DB 360 AKPEGSSASRPVPASRGSKTLCKGDRQAPGPPARFPRPIWSASPPAPRSSTPCPGA 419
QY 421 VREDTYPVGTGCVSPALAOGGPOGSMRFLQMSNMPRLPTDLVVGPMFRHDFRQSCMV 480
DB 420 VREDTYPVGTGCVSPALAOGGPOGSMRFLQMSNMPRLPTDLVVGPMFRHDFRQSCMV 479
QY 481 RAISOEDQALPCWQAEHPAE 500
DB 480 RAISOEDQALPCWQAEHCGE 499
```

RESULT 9

US-10-108-260A-4446
; Sequence 4446, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108.260A
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4446
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4446

Query Match 14.5%; Score 433.5; DB 15; Length 210;
Best Local Similarity 53.3%; Pred. No. 1e-25;
Matches 89; Conservative 26; Mismatches 47; Indels 5; Gaps 2;

```
QY 18 IIMKEKGRAGLPEDKGPCK--PRSYNNNVHIGIYHETELPPLTAREAKQIRREISRK 75
DB 19 IITKEGGRAGLAVDLGBOVDVAKTYNN--LGIYHEMELPRVSALEVKORRESKRT 75
QY 76 SKVYDMLGDMKEXKSSRLIDRAYKGMPMNIRGPMVSVLNIEEMKLNQPRYQIMKEKG 135
DB 76 NKWQMLADWTKYTKSLQGVCKVIPLAVRGALSLDLIDIKSKQNPQKVKYKMEKG 135
QY 136 KRSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEERNP 182
DB 136 KRSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEERNP 182
QY 136 KRSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEERNP 182
DB 136 KRSEHIQRIDRVSGTLRKHIIFPRDRYGTQORELHILAYEERNP 182
```

```
RESULT 10
US-10-425-115-237632
; Sequence 237632, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT FILING DATE: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 237632
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: MRT4577_148305C.1.pep
US-10-425-115-237632

Query Match
Best Local Similarity 11.7%; Score 348.5; DB 17; Length 381;
Matches 117; Conservative 54; Mismatches 161; Indels 63; Gaps 13;

QY 33 DKGPFRSYNNNVHDLGI-----VHETPLPPLTAREAKQIRREISRK 75
DB 11 EEPGAPBSRA---VDRFGIKPEQGNSPDGIPIKKSIIHRE-----REERRIR----- 54
QY 76 SKVNDMLG---DWEKY--KSSRKLIIDRAYKGMPMNIRGPMWSVLTINIEEMKLNPERYQ 129
DB 55 -KMRKMGIVGNDMWHYRRNPHVYKRIKGIIPDCIRGLWOLISGRDLMLNPGVYE 113
QY 130 IMKEKGRKSSHHQRIIDVSGTLRKHIIPDRYGTQRELIHLLAYEENPEVGYCD 189
DB 114 TLVIYETSASB--LEIIRDISRTPSHIFQQRHGPQGRSLVNLKAYSVDYDVGYYG 171
QY 190 LSHIALFLIYLPREDAFMALVOLASERHS-LOGFHSPPGCTVQGLQDQOEHVATSP 248
DB 172 MGFLAGLLLTMSBEDAFMLLVALLKGAHAPMEGLYQAGLPVQOYLSQFEKLVIELMP 231
QY 249 KTMGHODKDLCCGQCSPLGCLIRILIDGSLGLTLELMVYLVESGQALMPITRIAFKYQ 308
DB 232 KLQGHFVEEMINPSMVASQWFIIVFSYSPFFHLTLKVMDFLYEGIKVYFQVGLALLRLC 291
QY 309 QKRLTSTRSGPMPAR--FCNRFPVDTWARDDETVLKLIRASMKKLTIRKQGDLEPPAKPEQG 366
DB 292 HDLVLK-----PPEKLLYALRNPEEPATDPD-VLPLAFSPVSSRLBELQKEYQKQEG 346
QY 367 ----SSASRPVP-----ASRGKTLCKGDRO 388
DB 347 TSETSSGERLQPLVSKTMSRVGSRVSNLTTADRK 381

RESULT 11
US-10-425-114-64834
; Sequence 64834, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53131)B
; CURRENT FILING DATE: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
```

```
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64834
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: LIB4720-007-A6_FLI.pep
US-10-425-114-64834

Query Match
Best Local Similarity 11.7%; Score 348.5; DB 15; Length 434;
Matches 116; Conservative 56; Mismatches 157; Indels 57; Gaps 14;

QY 33 DKGPFRSYNNNVHDLGI-----VHETPLPPLTAREAKQIRREISRK 75
DB 64 EEPGAPBSRA---VDRFGIKPEQGNSPDGIPIKKSIIHRE-----REERRIR----- 107
QY 76 SKVNDMLG---DWEKY--KSSRKLIIDRAYKGMPMNIRGPMWSVLTINIEEMKLNPERYQ 129
DB 108 -KMRKMGIVGNDMWHYRRNPHVYKRIKGIIPDCIRGLWOLISGRDLMLNPGVYE 166
QY 130 IMKEKGRKSSHHQRIIDVSGTLRKHIIPDRYGTQRELIHLLAYEENPEVGYCD 189
DB 167 TLVIYETSASB--LEIIRDISRTPSHIFQQRHGPQGRSLVNLKAYSVDYDVGYYG 224
QY 190 LSHIALFLIYLPREDAFMALVOLASERHS-LOGFHSPPGCTVQGLQDQOEHVATSP 248
DB 225 MGFLAGLLLTMSBEDAFMLLVALLKGAHAPMEGLYQAGLPVQOYLSQFEKLVIELMP 284
QY 249 KTMGHODKDLCCGQCSPLGCLIRILIDGSLGLTLELMVYLVESGQALMPITRIAFKYQ 308
DB 285 KLQGHFVEEMINPSMVASQWFIIVFSYSPFFHLTLKVMDFLYEGIKVYFQVGLALLRLC 344
QY 309 QKRLTSTRSGPMPAR--FCNRFPVDTWARDDETVLKLIRASMKKLTIRKQGDLEPPAKPEQG 366
DB 345 HDLVLK-----PPEKLLYALRNPEEPATDPD-VLPLAFSPVSSRLBELQKEYQKQEG 399
QY 367 ----SSASRPVPASRGKTLCK--GDR 387
DB 400 TSETSSGERLQPLA--SKTMSRVGSR 423

RESULT 12
US-10-425-114-70559
; Sequence 70559, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53131)B
; CURRENT FILING DATE: US/10/425,114
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70559
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; OTHER INFORMATION: Clone ID: UC-ZMFLBOSINTB111D05_FLI.pep
US-10-425-114-70559

Query Match
Best Local Similarity 11.7%; Score 348.5; DB 15; Length 434;
Matches 117; Conservative 54; Mismatches 161; Indels 63; Gaps 13;

QY 33 DKGPFRSYNNNVHDLGI-----VHETPLPPLTAREAKQIRREISRK 75
```

```
Db      64 EFGPAPSR-----VDRFGFIKEQGNSPDGIPEKGSIHRE-----REERRIR----- 107
Qy      76 SKVVMIG-----DMEKY--KSSRKLIIDRAYKGMPPNINRGPMWSVLLNTEEMLKXPRGYQ 129
Db      108 -KWRMTIGGNDMGHYRARNPHVVKRRIRKGIPOCLRGLVMQLISGRDLILMPGYVE 166
Qy      130 IMKEKGRSSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQORELHIIILAYEYEVGYCSD 189
Db      167 TLVIYETASB--LEIINDISRTFSPSHIFQGRHGRGQSLNLIKAYSVIRDRDGYVGG 224
Qy      190 LSHNAPLFLYIPEDDAFWALVOLLSERHS--LOGFHSNGGTVOGLDQOEHVAVTSOP 248
Db      225 MGFLDLQLLYMSEBDAFWLLVALLKGAHVAMEGLYQGLPLVOQYLSQFEKLYIELMP 284
Qy      249 KTMGHODKDKCGGCSPLGCLIRILIDGSLGLTLRLMPVYVEGEQALMPTRIAFKYQ 308
Db      285 KLGQFVEMINPSYASQWFTVFSYSPFPFLTLRWMDVFLYEGIKVFOVGLALLRQC 344
Qy      309 QKRLTKSRGCFMAR--FCNRFVDTWARDDETVLKLHRSMMKLTTRKQGLDPPAKPEQG 366
Db      345 HDDLVKL-----PFEKLVALARFPEEATDPD--VLPLAFSPKVSRLBELQKRYQKRGQ 399
Qy      367 ---SSASRPVP-----ASRGKTLCKGDRQ 388
Db      400 TSETSSGERLOPLVSKTMSRVGSRVSNLTTADRK 434

RESULT 13
US-10-437-963-139162
; Sequence 139162, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Mu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 139162
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Oryza sativa
; OTHER INFORMATION: Clone ID: PAT_MFT4530_40481C.1 Dep
US-10-437-963-139162

Query Match      11.0%; Score 329; DB 16; Length 433;
Best Local Similarity 29.8%; Pred. No. 3.9e-17;
Matches 98; Conservative 45; Mismatches 126; Indels 58; Gaps 9;

Qy      33 DKGPAPFSSYNNVNDHLG-----IVHETB-----LPLPLTARE 64
Db      11 EFGPAPSR-----VDRFGFIKEQGTSPDGISKSFHIBREBSGYWLTVHFTLADI 66
Qy      65 AKQIRREISRSKSYVDNG-----DMEKY--KSSRKLIIDRAYKGMPPNINRGPMWSVLLNTE 118
Db      67 FLGNHRDRRIKKMKMIGVGSGDWKHYRARNPHVVKRRIRKGIPOCLRGLVMQLISGR 126
Qy      119 EMKLKPNRYOIMKEKGRSSSEHIQRIIDRVSGTLRKHIIFPRDRYGTQORELHIIILAYE 178
Db      127 DILLNPFVETLVYETASB--LEIINDISRTFSPSHIFQGRHGRGQSLNLIKAYSVIRDRD 184
Qy      179 EYNBEVGYCRDLISHIALFLYLPEDDAFWALVOLLSERHS--LOGFHSNGGTVOGLD 237
Db      185 VYDRDVGTVQMGFLAGLLLYMSEBDAFWLLVALLKGAHVAMEGLYQAGLPLVOQYLY 244
```

```
Qy      238 OOEHVAVTSOPKTMGHODKDKCGCSPGLCLIRILIDGSLGLTLRLMDVYLVGEQAL 297
Db      245 QPEKLVLMQMPOLGHPFIEEMINPSYASQWFTVFSYSPFPFLTLRWMDVFLYEG----- 300
Qy      298 MPETRIAFKVOQKRLTKTSRCGPMARFCN 326
Db      301 ---IKVFOVGLALL-----RFCH 316

RESULT 14
US-10-051-874-138
; Sequence 138, Application US/10051874
; Publication No. US2004000557A1
; GENERAL INFORMATION:
; APPLICANT: Pedigaru, Muralidhara
; APPLICANT: Alisbrook II, John P
; APPLICANT: Coleman, Steven D
; APPLICANT: Spylek, Kimberly A
; APPLICANT: Boldog, Ferenc
; APPLICANT: Vernet, Corine AM
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Casman, Stacie J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Patutarajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Pena, Carol BA
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Zetnusen, Bryan D
; APPLICANT: Miller, Isabelle
; APPLICANT: Miller, Charles B
; APPLICANT: Lepley, Denise M
; APPLICANT: Smithson, Glenda
; APPLICANT: Baumgartner, Jason C
; APPLICANT: Heitman, John L
; APPLICANT: Peyman, John A
; APPLICANT: Gorman, Linda
; APPLICANT: Mezes, Peter D
; APPLICANT: Kaupia, Ramesh
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grose, William M
; APPLICANT: Liu, Xiaohong
; APPLICANT: Ellerman, Karen
; APPLICANT: Rothenberg, Mark
; APPLICANT: Stone, David J
; APPLICANT: Bugess, Catherine E
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE OF INVENTION: USING THE SAME
; FILE REFERENCE: 21402-245
; CURRENT APPLICATION NUMBER: US/10/051,874
; CURRENT FILING DATE: 2002-09-25
; PRIOR APPLICATION NUMBER: 60/268,595
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/325,306
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/262,587
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/272,409
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/262,454
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/276,777
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/291,672
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/330,336
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/265,530
```


Qy 283 -----LRLMDVYLVEGEOALMPITRIAFKVOOKRLTKTSRCGPMARFCNRFVDTWA 333
 Db 254 ARTLPMASVLRWMDMFPCBG-----VKIIFRVALVLRHTLGSVEKLRSCOGMYETWE 306
 Qy 334 R-----DEDTVLKHL-----RASMKKLTRKQGD----- 357
 Db 307 QLRNLPQOCMEDEFLVHEVTNLPTVEALIERENAAQLKKMRETRGELQYRPSRRLHGSRA 366
 Qy 358 -----PPPAKPPQGSASRPVPASRGCKTLCKGDRQAPPGPPARPPRPIMSASPPRA 409
 Db 367 IHEERRROQPPILGP-SSSLSLPLGLKSRGSR--AAGAPSP-PPVR-----RA 411
 Qy 410 PRSSTPCPGAV-REDTYPVGTQGVSPALAOGPGS 446
 Db 412 --SAGPAPGPVVTAGLHP---SLPSPT-GNSTPLGS 442

Search completed: February 4, 2005, 08:06:45
 Job time : 88 secs

This Page Blank (uspio)